CITY UNIVERSITY LONDON

RISK AND REGULATION OF ACCESS TO PERSONAL DATA ON ONLINE SOCIAL NETWORKING SERVICES IN THE UK

JOHN DAVID HAYNES

DOCTOR OF PHILOSOPHY

CITY UNIVERSITY, LONDON

CENTRE FOR INFORMATION SCIENCE

SCHOOL OF MATHEMATICS, COMPUTER SCIENCE AND ENGINEERING

APRIL 2015

ACKNOWLEDGEMENTS

I am immensely grateful to my parents, Patrick and Ruth Haynes, for an amazing childhood and a great start in life. I have also been astonishingly lucky in my two supervisors David Bawden and Lyn Robinson who gave me the latitude to make my own way, but who were always there to provide guidance in the most difficult of times. My thanks to Tamara Eisenschitz formerly of City University and Alan Bell at the University of Dundee who planted the idea in my mind of doing academic research. Members of staff at City University Library were invaluable in helping me to find my way around the law literature. Lunches, seminars, coffee meetings with friends and colleagues in the School of Mathematics, Computer Science and Engineering and the School of Law at City University gave me the motivation to come in to University each day. Finally, I pay tribute to William Blacklock, my late partner, without whose encouragement and support I would never have started this enterprise.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

This research investigates the relative effectiveness of different modes of regulation of access to personal data on social networking services in the UK. A review of the literature demonstrated that there was a gap in research comparing different regulatory modes applied to online social networking services (SNSs). A model of regulation was developed based on Lessig's four modes of regulating the internet. Risk to individual users was selected as a way of testing different regulatory approaches, using the premise that risk-based regulation has become a key consideration in European regulation. The regulatory effects were tested using: online surveys, interviews with industry experts, content analysis of privacy policies, and a legislative review. The research data are appended to the main body of the thesis. The research demonstrated the potential of risk as a means of distinguishing between different regulatory modes and concluded that a combination of regulatory approaches was the most effective way of protecting individuals against abuse of personal data on online SNSs. Further research suggested includes: looking at risk from the perspective of companies, and of society; further development of the regulatory model; and country comparisons to discover whether the findings of this study are more generally applicable.

SECTION I. INTRODUCTION

CHAPTER 1 – INTRODUCTION

BACKGROUND

The idea for this research arose from the personal experience of the researcher as an information consultant working extensively in the public, private and voluntary sectors. By 2010, when this research started, information governance had become a major concern for many organizations and legislation such as the Data Protection Act 1998 prompted a strong focus on compliance issues.

From 2004 onwards first large-scale online social networking services (SNSs) began to take hold. By 2010 there were increasing concerns about privacy and abuse of personal data. As companies began to consider the benefits of using social media for marketing there was widening awareness of the risks posed by the disclosure of personal data. This preoccupation with risk presented an opportunity to consider ways in which access to personal data is regulated.

Although social media are global in nature, they are delivered to users with an expectation that they are protected by national laws. This study considers SNSs available to UK users and examines the regulatory landscape that exists in the UK.

An example of the concerns that initiated this study is the reported declaration by Mark Zuckerberg that the age of privacy was over (Johnson 2010). In 2010 the default privacy settings for Facebook's then 310 million users were changed to make their profiles public. This caused an outcry and forced the company to revert to a private default, and to update its privacy policy.

Facebook, as a US-based company, is subject to the voluntary, self-regulatory U.S.-EU Safe Harbor framework run by the US Federal Trade Commission (International Trade Administration 2009). It is also subject to public pressure and this could be argued to be a form of regulation directly influenced by users. Are these two methods of regulation effective means of protecting personal data, or are there more effective methods that could be applied? For instance, does the legislative regulatory approach adopted in the UK with the Data Protection Act 1998 afford better protection of personal data?

One way of testing this is to compare the ways in which social media, and specifically online social networking services (SNSs) respond to data protection regulation in the UK. There may be variations in attitudes to regulation and the different modes of regulation that may apply to this sector to protect users against misuse of their personal data. For instance, SNS providers

who are based in the United States, and are members of the self-regulating U.S.-EU Safe Harbor scheme have adopted different provisions of the scheme (Connolly 2008). Does this difference in approach make a material difference to UK users and are they exposed to greater risk as a result of non-compliance with EU legislation?

WHAT IS AN ONLINE SNS?

Before going any further it helps to outline what is meant by online Social Networking Services (SNSs) and why regulation of access to personal data might be significant and topical. Online SNSs are internet services based on individual members who put up profiles (containing personal information) that are available to other users or members of the service. Users are able to link to other members to build up their own personal networks. This may include concepts such as: 'linking'; 'connecting'; 'following'; and 'friending'. In some cases personal profile information may be limited to authorised members who have been specifically identified by a user as being part of their network. In other instances profiles may be available to all users of the service. At the most extreme end of this range, the information in personal profiles may be available to general internet users regardless of membership of the network service. The area of debate is around the release of personal data by the social network providers to third party external agencies (such as advertisers and recruitment agencies) so that they can target their marketing. Tracking technologies such as cookies and beacons are widely used by social network providers to pass on detailed information to commercial enterprises for behavioural advertising.

In a series of articles the Wall Street Journal described the tracking technologies used by the most popular websites in the United States, including many of the largest SNSs (Angwin & Mcginty 2010). Cookies are the most widely used tracker, and have been the subject of recent European legislation, implemented in the United Kingdom as the Privacy and Electronic Communications (EC Directive) (Amendment) Regulation (SI 2011/1208). A cookie is a small text file with a unique identifier that the website places in the user's browser to keep track of the online session. Many websites require cookies to operate effectively (they ensure the continuity of a session and can be useful if there is an interruption or if a user logs off and then logs back onto a website).

Where a website provider belongs to an advertising network or purchases the services of a tracking company, a third-party cookie tracks usage from site to site and records this data centrally for exploitation, by selling to advertisers, for instance. These advertisers, without necessarily knowing very much about an individual user, can target advertisements to them

automatically. For instance, a search of a travel site may indicate interest in visiting a particular city and subsequently ads for hotels in that city may start to appear in banner ads when browsing. This technology is discussed in greater detail in Chapter 9.

Gathering all this personal data allows for extensive data mining by data aggregators. An alternative approach has been taken by Blue Cava, which has developed a very large database of all the devices on the Internet (Auerbach 2013). It is constructed by creating a unique fingerprint of each device based on its settings, which are detected by web servers. Once a device has been identified it is possible to track online behaviour and to generate massive data sets that can be exploited by data mining techniques to target services and advertisements. Another technology increasingly used is location-based tracking of mobile devices such as smart phones and data pads.

Sites without adequate security can be subject to scraping where an anonymous login allows access to the profiles of other members of the site, thus enabling the scraper to capture personal data without reciprocating. Linking data gathered online with offline and published personal records (such as post codes or electoral registers) or blogs and tweets is also of some concern (Information Commissioner's Office 2012).

Research Objectives

This research aims to analyse the regulatory landscape and to compare the advantages and disadvantages of different modes of regulation. It uses a model of regulation based on an analysis of formal and informal measures to protect users from misuse of personal data. The research combines risk analysis and a legislative review to assess the relative effectiveness of different modes of regulation and to gain an insight into ways in which the regulatory landscape might be changed to protect users more effectively.

This work sets out to test the hypothesis that **law-based regulation alone is not the most effective way of protecting users against the risks associated with use of SNSs**. It also considers the hypothesis that **risk to individual users can be used to compare the effectiveness of different modes of regulation**. Risk assessments are used to compare the different modes of regulation. In order to test these hypotheses a number of research questions need to be asked:

- What is the nature of regulation of access to personal data on social networking services (SNSs)?
- What are the risks to users of having personal data on SNSs?

- How have law-makers responded to the inception and growth of SNSs?
- What effects do different regulatory methods have on risk to individuals?
- Is risk to users an effective method of comparing different modes of regulation?

THESIS STRUCTURE

This thesis is divided into four main sections with individual chapters arranged by theme within each section:

Section I. INTRODUCTION

Section II. RISK

Section III. REGULATION

Section IV. CONCLUSION

In Section I the research questions are introduced in Chapter 1 and prior research in the area is investigated by means of a literature review in Chapter 2. Chapter 3 describes the methodology and in particular the development of a risk-based approach to evaluation of regulatory modes.

Section II considers what personal data is and works toward a definition of information privacy in Chapter 4. Chapter 5 defines risk and develops a model of personal risk that can be used to compare the effects of different modes of regulation.

Section III considers different approaches to modelling regulation before developing a model based on Lessig's modes of internet regulation (Chapter 6). Each of the new model's modes of regulation is described and investigated in turn: Legislation (Chapter 7), Self-regulation through privacy policies (Chapter 8), Self-regulation of digital advertising (Chapter 9), Code (Chapter 10) and Norms (Chapter 11).

In Section IV the thesis revisits the research questions and discusses the results of the different research investigations (Chapter 12) before pulling together all the strands in the Conclusion (Chapter 13).

The appendices contain details of the research instruments and data from the investigations that support the arguments put forward in this thesis and the attached CD-ROM contains data sets gathered during this research.

CHAPTER 2 – LITERATURE REVIEW

REVIEW METHODOLOGY

Preliminary reading focused on the key texts in the areas of: regulation, research methodology, and relevant legislation (Room 2007; Jay 2003; Baldwin et al. 2012; Pickard 2013; Brown & Marsden 2013). Some of these texts were identified from academic departmental reading lists, and searches of library catalogues (notably the M25 union catalogue for academic and research libraries in and around London, and the British Library).

Literature from peer-reviewed journals was identified by searching a range of aggregated electronic journals and bibliographic databases. They were chosen for their subject emphasis, comprehensiveness and the features that they offered:

GoogleScholar – general bibliographic database resource used as a starting point and as a 'catch-all' to supplement searches on more specialist systems

EBSCOHost – aggregated journals

Emerald – aggregated journals

World of Knowledge – social science and science bibliographic database including Social Science Citation Index

LexisNexis – comprehensive legal journals and news reports

ACM – informatics journals

UK and European legislation from www.legislation.gov.uk and http://eur-lex.europe.eu.

Wikipedia and Google Search were also used as sources to supplement the scholarly services above. Items identified through these two routes were evaluated in the same way as those from the scholarly sources.

Key papers were identified by sorting on citation count coupled with an analysis of recent downloads for newer papers that may not have had time to build up a significant citation count.

Bibliographic search results were assessed for relevance by title and abstract. Where there was doubt, an assessment was based on examining the full text. The majority of journal articles cited in this review were downloaded as e-journal articles. Most books were read as hardcopies at the British Library, although in a few cases e-book versions were available and were downloaded.

Although this was a systematic literature review, as defined by Bawden and Robinson (2012, p.316), some sources of information were identified by following references made by speakers at seminars and conferences, in blogs, on websites and via social media services. References identified in this way were tracked back to the original source and retrieved from peer reviewed journals or formal publications where possible.

The original literature review was conducted in 2012 during preparation of the transfer report. After that point, as new sources became available they were incorporated into the Mendeley database compiled for this research. The literature review was updated in November 2014 to incorporate material published from 2012 onwards.

Search Strategies

This research sets out to test the hypothesis that assessing risk to individuals can be used to compare the effectiveness of different modes of regulating access to personal data on SNSs.

This is a large and complex question and for the purposes of a literature review can be broken down into a series of search queries. This approach allows more comprehensive searching of the component concepts in the main questions and allows navigation and refinement. The expectation was that the research question had not been expressed in this form before and that there would therefore be little literature specifically addressing the question. The focus of the literature review was on the following questions:

- What are the risks associated with making personal data available via social networks?
- How is access to personal data on social networks regulated?
- How is the relative effectiveness of different modes of regulation evaluated?

Embedded within these questions are a number of concepts that require further exploration. These are addressed in the discussion below, which considers each of the review questions in turn.

What are the Risks?

The search is broken down into three concept areas as seen in Figure 1. Within each concept are synonyms that express that concept. In some cases truncation was necessary to spread the search scope. The process was iterative and the three concepts were combined using the Boolean AND operator to narrow the search down successively. The first concept was enhanced by including the name of the two most frequently cited SNSs in the literature. An analysis of search performance using EBSCOhost using the OR operator to link the terms

"Social Network*", "Google" and "Facebook" increased the retrieval of scholarly articles since 2012 by 45% (from 8,732 to 12,662 hits). Further searches with this group of terms showed that inclusion of Ning provided an additional 3% increase (from 12,662 to 13,043 hits) and Twitter an additional 2% increase (from 12,662 to 12,958 hits) (Appendix A). Examination of the additional hits found no relevant articles. The remaining SNSs increased the hit rate by less than 0.3%. Table 23 in Appendix A contains the results of this sensitivity test.

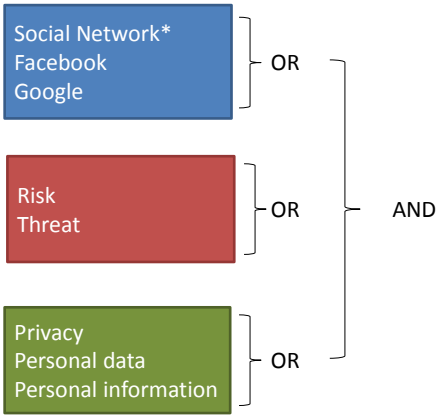

FIGURE 1 - SEARCH QUERY ON 'RISKS' AND 'SOCIAL NETWORKS'

The core theme for this search was online social networking services. A general search (which will pick up some irrelevant material – such as 'offline' social networks) was narrowed down by introducing the concepts of risk and privacy. A manual trawl of the resulting items eliminated articles that were not about the risks associated with personal data – for instance reputational risk to organisations where an employee defames a customer or a supplier or competitor using social media.

HOW IS ACCESS TO PERSONAL DATA ON SOCIAL NETWORKS REGULATED?

The next search, Figure 2, combined the concepts of 'social network' with 'regulation'. This was narrowed down specifically to consider privacy and personal information to eliminate literature on copyright.

A key element of this research was to identify the different modes of regulation of access to personal data, with reference to social networks. For the purposes of this question 'personal data' is taken to mean data relating to an individual that can be uniquely associated with that individual. It refers to data that would not normally be in the public domain such as their political or religious views, information about their personal or private life, their personal finances and their behaviour or habits.

Personal data includes aggregated data as well as identifying data that can be associated with a specific individual, whether named or not.

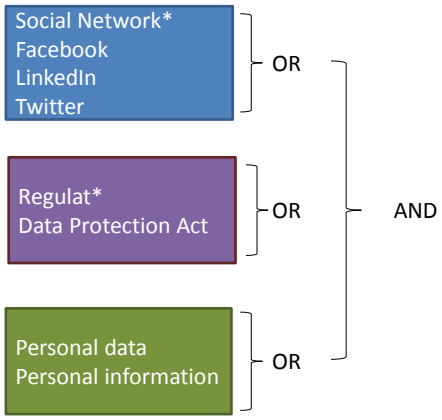

FIGURE 2 - SEARCH QUERY ON 'REGULATION OF PERSONAL DATA ON SOCIAL NETWORKS'

HOW IS THE RELATIVE EFFECTIVENESS OF DIFFERENT MODES OF REGULATION EVALUATED?

The third query, Figure 3, combines the concept of effectiveness/evaluation with regulation. There is a large, well-established body of literature on regulatory effectiveness and evaluation. The terms around 'evaluation', 'effectiveness' and 'assessment' were combined using the Boolean 'OR' operator before being refined by the 'AND' operator for combining with the concept 'regulation'.

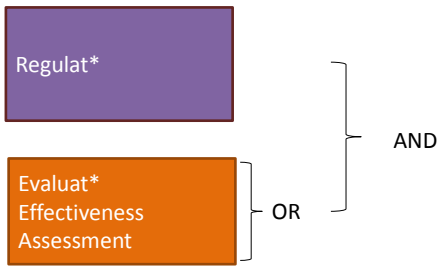

FIGURE 3 - SEARCH QUERY ON 'REGULATORY EFFECTIVENESS'

APPRAISAL CRITERIA

Where the search results were too broad, they were narrowed down to look at sources specifically relating to the UK, although care was taken not to eliminate literature about other territories that could provide comparative material for this study, or which might inform the methodology used for this study.

The scope of the searches focused on more recent material, mainly published from 2004 onwards, as this is when the current generation of SNSs emerged with the launch of Facebook. Apart from the paucity of literature about online SNSs before that date, this is a very rapidly evolving sector and older material is less relevant to the services available at the time of writing. Some exceptions were made for key references that are widely cited and which have had a strong influence on subsequent thinking, particularly in the area of regulation.

Bibliographic databases were selected for their subject coverage, with an emphasis on peer-reviewed journals. This was supplemented with manual searching of key journals identified during bibliographic searches and by recommendation and searches for key authors. Selected key papers were also searched on ISI's Science Citation Index to identify subsequent references. These searches were repeated periodically to update the literature review.

Primary and secondary legislation were also directly accessed and commentaries on the legislation were referred to. There was a strong emphasis on monographs and multi-author handbooks for the legal literature in addition to peer reviewed literature, published reports and academic studies.

CONTEXT FOR THIS RESEARCH

THE NATURE OF THE PROBLEM

This research considers the nature of risk associated with use of online SNSs and focuses specifically on risks associated with personal data on social networks. An internet user in effect makes a contract with an SNS provider when they sign up to a service. In exchange for providing personal data about themselves, they are given access to a range of services and features. They also benefit from having access to the personal profile of others on the network. This may be tightly controlled, so that they can only see details of individuals who have given them permission to view their profiles, for example a reciprocal arrangement like 'friending' on Facebook or 'Join my network' on LinkedIn. The less evident part of this equation is that in exchange for use of online SNSs, the service provider may sell on the personal data or access to personal behavioural data to third parties.

A comprehensive review of the research literature on Facebook considered five different areas of research: *"descriptive analysis of users, motivations for using Facebook, identity presentation, the role of Facebook in social interactions, and privacy and information disclosure"* (Wilson et al. 2012)

WHAT IS A SOCIAL NETWORK?

With the widespread recognition of the power of social media as a marketing and promotion tool there is a plethora of guides and manuals on how to exploit social media for business success available (Comm 2010; Clapperton 2009). Some of the more authoritative guides provide a starting point for a common understanding of what social media are and how social networks fit into this sector. For instance, Wollan, Smith and Zhou (2011) make the distinction between social media, which have been around for millennia and digital social media, which have grown with the internet and which correspond to current usage of the term 'social media'. Digital social media are described in terms of their characteristics:

- Peer-to-peer communications
- Content created and posted by users
- Easy to use
- Highly accessible, scalable and operates in real time
- Public and transparent

Online social networking services (SNSs) can be characterised by their use and there was a marked increase in their use between 2009 and 2011, mainly for social activities (89% of users) and with only 22% of users using it for informational activities (Dutton & Blank 2011).

Kaplan and Haenlein (2010) identify six types of social media which they go on to classify according to social presence/media richness on one scale and self-presentation/self-disclosure on the other. This starts to address one of the fundamental issues of control of personal data, which is explored later on.

TABLE 1 - SOCIAL PRESENCE V. SELF-DISCLOSURE (AFTER KAPLAN AND HAENLEIN, 2010)

	Low social presence/media richness	Medium social presence / media richness	High social presence / media richness
High self-presentation /self-disclosure	Blogs	Social networking sites (e.g. Facebook)	Virtual social worlds (e.g. Second Life)
Low self-presentation /self-disclosure	Collaborative projects (e.g. Wikipedia)	Content communities (e.g. YouTube)	Virtual game worlds (e.g. World of WarCraft)

Social networking services are placed firmly in the middle of the social presence / media richness scale and are classed as being high self-presentation / self-disclosure services.

Kaplan and Haenlein go on to describe social networking sites as "*applications that enable users to connect by creating personal information profiles, inviting friends and colleagues to have access to those profiles, and sending e-mails and instant messages between each other*."

Cavazza (2010) divides social media into categories based on functionality of the sites (with some overlap for services with multiple functionality):

> **Publish** – this is primarily for blogging, micro-blog, social stream services and wikis
> **Share** – allows users to share externally sourced media, including material that they have created themselves
> **Discuss** – bulletin boards and social search tools
> **Commerce** – including reviews of services and tools for e-purchasing
> **Location** – including event organisation and geo-location tools
> **Network** – personal and professional networks of contacts
> **Games** – social gaming, virtual worlds and casual gaming services.

This definition can be refined further by focusing specifically on social networking services which are a sub-set of social media. Boyd and Ellison (2007) make a distinction between social **networking** sites and social **network** sites. They define social networking sites (the subject of this research) as:

> … *web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system.*

The main online SNSs covered by this research (and conforming to the above definition) were selected on the basis of: their size, the fact that they offer a service in English, and their availability to UK users:

- Badoo
- Facebook
- Google
- hi5
- Instagram
- LinkedIn
- Myspace

- Ning
- Snapchat
- Twitter
- WhatsApp

(Wikipedia 2014)

EXPLOITATION OF PERSONAL DATA

The term 'personal data' can have a variety of meanings. Schneier (2010) suggests that different types of personal data on social networks will require different measures to protect them. He identifies the following categories of personal data (with some degree of overlap) that can be used as a typology of personal data:

> **Service data** is data provided to the service provider to set up an account
>
> **Disclosed data** is what users put up on their profiles, this may include content that they have created
>
> **Entrusted data** is like disclosed data, but is personal data provided for 'friends' contacts' sites on the social network service
>
> **Incidental data** is data that other people post about the user – which may include text, photos or moving images
>
> **Behavioural data** is data that the site collects about the user's behaviour on the internet, this may be restricted to the SNS, or may be based on monitoring his or her interactions with other web sites
>
> **Derived data** is that derived from other personal data to make assumptions about the user, his or her views or behaviour (or from a third party source such as a published directory)

A great deal of recent attention in the press has been devoted to behavioural data, which users may not have provided explicit, informed consent for the service provider to sell or use (Angwin & Mcginty 2010). The permission may be hidden in the terms of service that a user signs up to in order to gain access to the service. This information may then form the basis for intrusive advertising directed at the user when they open their browser.

Warren and Brandeis (1890) provide a starting point for a definition of privacy which is discussed in detail in Chapter 4.

Discussions of privacy are made more complicated by the differing definitions of privacy. Krotoczynski (2013) argues that there is a fundamental difference in concept between the United States with its emphasis on freedom of speech and Europe with its emphasis on preservation of human rights values such as dignity and privacy. Unless this difference in concept and meaning is acknowledged it is not possible to implement global human rights values as expressed in the Universal Declaration of Human Rights 1950. This is a particularly important consideration when it comes to extra-territorial jurisdiction.

Wacks (2010) contends that there is no right of privacy by statute in UK and that protections arise from common law. That said, it could be argued that privacy is provided for in a number of UK General Public Acts including: the Data Protection Act 1998, the Communications Act 2003 and the Human Rights Act 1998. These are discussed in more detail in Chapter 7.

Bélanger and Crossler (2011) in their review of privacy research contend that privacy research in the context of information systems tends to be very US-centric and often focuses on student surveys. This limits their generalizability and they recommend that *"Researchers should utilize a broader diversity of sample populations"*. Their meta-analysis of the literature grouped theory types into the following categories according to Gregor's (2006) taxonomy:

- Analysing theories
- Explaining theories
- Predicting theories
- Explaining and predicting theories
- Design and action theories

Bélanger and Crossler also recommended that *"Researchers should conduct more studies investigating the why related to privacy as opposed to the how"*.

Some commentators have distinguished between 'information privacy' and wider definitions of privacy. For example Westin (2003) includes in his definition of privacy: "*the claim of an individual to determine what information about himself or herself should be known to others*". This is a theme comprehensively addressed in Smith, Dinev and Xu's (2011) review of Information Privacy Research. They make the distinction between physical privacy, which "*concerns physical access to an individual and/or the individual's surroundings and private space*" and information privacy, which "*concerns access to individually identifiable personal*

information." In their analysis of research the authors suggest that much privacy research has measured privacy concerns "*because of the near impossibility of measuring privacy itself*". Privacy concerns include beliefs, attitudes and perceptions about privacy. This in turn may be informed by the experiences of individuals, their awareness of privacy issues, personality differences, as well as demographic differences and the general culture or climate within which they live. In their APCO (Antecedents, Privacy Concerns, Outcomes) model Smith et al point out the current privacy research rarely considers outcomes as a result of privacy concerns.

Gurses (2014) in considering whether privacy can be engineered provides a thoughtful overview of privacy definitions, which is useful when considering options for technological approaches to privacy regulation. Solove (2007) goes on to discuss the common problem of conflating 'privacy' and 'secrecy', the so-called 'secrecy paradigm', and suggests that these two concepts need to be distinguished from one another.

The OxIS 2011 survey, found that nearly half of those surveyed thought that "*the current use of the internet is a threat to privacy*" (Dutton & Blank 2011). However a large number of respondents were happier to give out sensitive personal data. Next Generation Users[1] were less likely (42%) than non-users and ex-users (63%) to consider the use of computers and the internet a threat to privacy.

For gathering personal data Christiansen (2011) distinguishes between "*a user's voluntary sharing of such information*" and "*involuntary/uninformed collection by other parties*". She goes on to describe three types of data collection:

> *1. Collect personal data, then anonymize and aggregate it to sell to third parties and/or for use internally*
>
> *2. Collect personal data, keeping personal data within the company but providing the opportunity for advertisers to specify a certain range of traits for target marketing*

[1] Next Generation User "...someone who accesses the internet from multiple locations and devices. ...someone who uses at least two internet applications ... on their mobile or who fits two or more of the following criteria: they own a tablet, own a reader, own three or more computers." (Dutton and Blank, 2011, p4)

3. Collect personal data with the intention of selling the information, sometimes including specific profiles or names, to third parties

Christiansen (2011) also identifies some of the ways in which personal data can be used or misused, although this is far from comprehensive:

- *Running of background checks by employers for hiring decisions;*
- *Pricing and assessing risk of injury or death by insurance companies (based on Internet searches, blogs, and confidential online support groups, for example);*
- *Termination decisions by employers (for example, if a user is criticizing the workplace or found to be lying to the employer);*
- *Recruiting and scholarship decisions by athletic coaches;*
- *Searching for relevant evidence by attorneys in the course of case preparation;*
- *Detecting political leanings for fundraising purposes and to target individuals who are undecided on an issue or a candidate; and*
- *Facilitating criminal attacks.*

These can be addressed by legislation, by means of 'do not track' lists (rather like the telephone preference service) and ad blockers in web browsers. Finally Christiansen advocates better public education as a key means of reducing the risks associated with use of social media.

McGoldrick (2013) suggests that the European Court of Human Rights ruling in *Google Spain SL, Google Inc. v AEDP, Mario Costeja González* about the 'right to be forgotten' represents a reassertion of the societal value of privacy, which needs to be protected.

RISK

The focus on risk is a key element of this research. The English word 'risk' is derived from the Italian verb 'riscare' to run into danger. It acquired its commercial meaning in the 18th century when it was applied to insurance losses. Since then, the usage of the word 'risk' has moved away from the idea of a measurable hazard calculated on probability and size of loss, to

a more nebulous range of uses encompassing more subjective assessments of likelihood and impact. It is applied extensively in project management and in general management of organisations as well as in more traditional commercial environments such as banking.

In a UNESCO workshop Mansell (2008, p.16) identified risk as a significant area of potential research in information and communication studies. The report stated:

> *Research is needed on the potential of digital communication and information networks to produce data about all human activity and with respect to all people.*
>
> - *How can we evaluate the impact of the development of computing and artificial intelligence needed to engage in surveillance?*
> - *What is the potential for the (mis-)use of such data by government institutions or private enterprises?*
> - *What are the dangers or risks to the public interest?"*

Fischhoff et al (1984, p.124) suggest that it is often difficult to be objective in the assessment of risk:

> *Within the philosophy of science, 'objective' typically means something akin to 'independent of observer'. That is, any individual following the same procedure should reach the same conclusion. However meritorious as a goal, this sort of objectivity can rarely be achieved. Particularly in complex, novel areas, such as risk analysis, research requires the exercise of judgement.*

They go on to say: "*Thus, objectivity should always be an aspiration, but can never be an achievement of science*." They argue that researchers need to state what dimensions of risk they are considering: "*an analysis of 'risk' needs to specify which of these dimensions will be included. In general, definitions based on a single dimension will favour technologies that do their harm in a variety of ways (as opposed to those that create a lot of one kind of problem)*." In other words the way of measuring each dimension of risk will have an effect on the overall assessment of risk and as they state: "*Evaluating it fairly requires knowing what it was intended to accomplish.*" The steps needed to do this, they suggest, are: to decide which consequences to include; the development of a risk index based on different risk attributes; identification; and application of simplifying assumptions to make the problem 'tractable'.

The different attributes for a technology make up a vector which can then be turned into a single number. Fischhoff et al (1984, p.137) argue that "*Developing a definition of risk requires a variety of explicit value judgments*."

Aven and Renn (2009, p.2) review a number of definitions of risk before arriving at a new definition of risk that addresses some of the concerns they raise about lack of precision and the need to separate risk measurement and the decisions on the response to risk:

> *Risk refers to uncertainty about and severity of the events and consequences (or outcomes) of an activity with respect to something that humans value.*

They point out that:

> *The assigned probability and judgements about risk tolerability or acceptability are different dimensions, or separate domains in the world of risk professionals who make a clear distinction between risk assessment and judgement of acceptability.*

This allows for a qualitative approach to risks where probabilities or outcomes cannot be quantified. They conclude by stating that:

> *Our proposed definition provides a conceptually consistent and practically compatible concept of risk without falling into the extreme of total subjectivism and relativism but also not pretending that risk is a measurable object similar to other physical entities.*

Objectivity or subjectivity is a theme taken up by Hansson (2010, p.237), who concludes that "*risk is both fact-laden and value-laden and that it contains both objective and subjective components.*" He suggests that:

> *The real challenge is to identify the various types of factual and valuational components inherent in statements about risk and to understand how they are combined.*

There is a considerable body of research into risk and the internet (Nissenbaum 2010; Desai, MonicaLodge, DeborahGates, MelodiWolvin, MariaLouer 2012) and significant studies on risk associated with SNSs (Rosenblum 2007; Toch et al. 2012; Kiser 2011; Sellars 2011). Much of this focuses specifically on risks to children – an area that raises specific issues about

vulnerability and this is beyond the scope of this study (Slavtcheva-Petkova et al. 2015; Staksrud & Livingstone 2009; Livingstone 2013). The focus of this research is on risks to adults that are associated with use of SNSs and on ways of measuring those risks.

The Oxford Internet Institute (OII) is conducting a longitudinal survey OxIS which regularly surveys public attitudes to the internet and perceptions of risk. An OxIS report describes the paradox of increases in bad experiences on the internet and an increasing trust in the internet (Blank 2010). In other words bad experiences do not seem to undermine trust in the internet. It is suggested that the most experienced users of the internet are more likely to have had a bad experience but are better able to cope with problems that subsequently arise, because they are more experienced.

Xu, Michael and Chen (2013) have developed a model of 'privacy disclosure' (i.e. disclosure of personal information on SNSs) which is affected by perceptions of privacy risk and information control. This is balanced against perceived benefits such as 'social capital' and 'community attachment', which tend to motivate greater disclosure. Curiously they found that benefits such as access to free or reduced cost services were less of a motivating factor than the social factors. The research does not take into account the personal characteristics of individual users, such as age and gender. Nosko et al (2012) for instance suggested that there is a difference in disclosure behaviour between women and men. Subsequent work also starting with the 'Theory of Planned Behaviour' by Saeri et al (2014) suggested that norms and perceptions of risk were significant factors affecting users' privacy behaviour online. They found contrary to a review of earlier research (O'Brien & Torres 2012) that trust was not a significant factor in determining user behaviour.

As well as explicit disclosure, there are risks associated with un-informed consent and inadvertent disclosure. For instance, social log-in represents a significant area of data leakage and some research has identified the risks associated with this and suggested technical solutions to limit this data leakage (Kontaxis et al. 2012). Carmagnola et al (2014, pp.194–195) in contrast looked at relationship data and non-personal information that can be associated with a personal profile and concluded:

> *In recent years, OSNs have been used by increasing numbers of people, with the result that much personal data are stored on these systems. Even if a social network's privacy settings ensure that user profile information is protected, they do not ensure protection from attackers who may combine disparate*

pieces of information about a user from multiple networks, thus allowing user identification and the retrieval of user personal data. Cross-site user identification may be very useful for the optimization of some tasks requiring user modelling, such as user support and personalization. However, it can also be used with criminal purposes, thus representing a risk for user privacy.

Weiss (2008) suggested that there needs to be a shift away from privacy-enhancing technologies that concentrate on access protection, anonymity and unlinkability. He advocates a move toward privacy safeguarding measures that enable greater transparency and that facilitate context and purpose limitation to personally identifiable data.

REGULATION

The following factors have affected the rise of regulation in the UK (Moran 2005):

- Crisis and scandal
- European Union - culture of greater regulation
- Power and democracy

One of the purposes of regulation is to control risk. In the introduction to the Better Regulation Commission's (2006, p.3) report, Rick Haythornthwaite, the Commission's Chairman, states:

Our specific recommendations are, though, for Government, and the most important of these calls is for our leaders to redefine our approach to risk management in a number of ways:

- *Emphasising the importance or resilience, self-reliance, freedom, innovation and a spirit of adventure in today's society;*
- *Leaving the responsibility for managing risk with those best placed to manage it and to embark on state regulation only where it represents the optimum solution for managing risk;*
- *Re-examining areas where the state has assumed more responsibility for people's lives than is healthy or desired; and*
- *Separating fact from emotion and emphasising the need to balance necessary levels of protection with preserving reasonable levels of risk.*

The BRC report (2006, p.51) goes on to identify five principles of good regulation:

- Proportionality "*regulators should only intervene when necessary. Remedies should be appropriate to the risk posed and costs identified and minimised*"
- Accountability "*regulators must be able to justify decisions and be subject to public scrutiny*"
- Consistency "*government rules and standards must be joined up and implemented fairly*"
- Transparency "*regulators should be open and keep regulations simple and user-friendly*"
- Targeting "*regulation should be focused on the problem and minimise side effects*".

These principles apply to government regulation and specifically regulation that is based on legislation. However some of these principles could be applied to other modes of regulation such as: self-regulation; and technology-based regulatory measures. Haythornthwaite (2006) suggests that government regulation should be based on a sound risk assessment and not be subject to public opinion, pressure groups and politics. However he recognises the need for a conversation between different interest groups in order to arrive at workable regulation. The administrative costs of regulation (red tape), is estimated at £30 – £40 billion a year in the UK. This is in addition to the cost of lost opportunities and non-financial costs such as: erosion of personal responsibility, loss of trust. This suggests that regulation should only be used where absolutely necessary. Haythornthwaite suggests that government should not regulate where the individual taking the risk is the only one that can be harmed. In other words, legislation may not have been considered the best way of protecting users against risk. In light of recent developments in the financial markets, the balance may shift back towards greater regulation (Heffernan 2011; Slattery & Nellis 2011).

REGULATING THE INTERNET

There are a variety of modes for regulating the internet generally and SNSs in particular. Ofcom (2009) identified the following types of regulation in relation to the internet:

- Personal responsibility – regulation by individual users
- Self regulation – regulation by site owners and content providers
- Self regulation – regulation by internet service providers
- Statutory regulation

An alternative view of regulation can be described in terms of a response to risk, and in particular risk associated with loss of privacy. Lessig (2006, pp.233, 234) identifies two main threats to privacy from the internet:

> *The first is the threat from 'digital surveillance' - the growing capacity of the government (among others) to 'spy' on your activities 'in public'* […] *The second threat comes from the increasing aggregation of data by private (among other) entities. These data are gathered not so much to 'spy' as to facilitate commerce.*

He identifies four modes of regulation, which he uses to mean "constraint" as responses to these risks:

> *Against these two different risks, we can imagine four types of responses, each mapping one of the modalities* […]*:*
> *Law* …
> *Norms* …
> *Markets* …
> *Architecture/Code* …

These models of regulation can be consolidated into a single model (discussed in Chapter 6), which form the basis for this research:

1. Legislation
2. Self-regulation
3. Code
4. Norms

LEGISLATION

A review of legislation and the resulting statutory regulation must take into account the nature of the UK constitution which governs the role of the different branches of government, including the legislature and the judiciary. Bogdanor (2009) argues that the current round of constitutional reform started in 1997 and the UK's entry into the Common Market in 1973 has effectively replaced one constitution with another. He also argues that this is an on-going process. Recent changes include: devolution; home country legislatures and the London Assembly; PR in devolved administrations; and European elections. Other significant developments that affect the regulation of access to personal data include: Human Rights Act

1998, reform of the House of Lords, Freedom of Information Act 2000, the regulation of political parties and electoral expenditure, and the Constitutional Reform Act 2005 which makes the Lord Chief Justice (not the Lord Chancellor) the head of the judiciary. Bogdanor argues that Britain has an uncodified constitution making it more difficult to follow changes to the constitutional order:

> *A society is distinguished from a mere conglomeration of individuals in that it comprises a group of people bound together by rules; and a constitution is nothing more than a collection of the most important rules prescribing the distribution of power between the institutions of government - legislature, executive and judiciary - and between the individuals and the state.*

He suggests that under the Human Rights Act 1998 judges are likely to play a more important role in defining rights (Bogdanor 2009). One purpose of a constitution is to protect minorities, which is what the Human Rights Act 1998 does. It is based on the European Convention on Human Rights and is part of the package of membership of the EU. Formerly sovereignty of parliament was the dominating principle in the UK's constitution; and hence a written constitution was deemed unnecessary. Now other factors are important and sovereignty of parliament is not the only consideration. Although Bogdanor argues in favour of a codified constitution, he advocates this as a process.

The Human Rights Act 1998 enshrines "*the right to respect for private and family life*". The other relevant piece of legislation is the Data Protection Act 1998 and Statutory Instruments under the Act and derived from the equivalent European Directives. These two pieces of legislation are described in Chapter 7.

Commentators such as Clarke (1999) advocated legislation in the US to protect personal privacy in the internet age. Oppenheim (2001) gives an early account of legislation relating to the internet and other electronic environments. In his chapter on Conflict of Laws he asks which law applies to a service based in one country but delivered to a user in another. He puts forward the idea that cyberspace should have its own jurisdiction and that its laws would apply to any transactions that took place in that space. Lessig's (2006, p.298) argument develops this theme to suggest that the norms and laws of both physical and cyberspace apply:

> *We have this desire to pick: We want to say that they are either in cyberspace or in real space. We have this desire because we want to know which space is*

> *responsible. Which space has jurisdiction over them? Which space rules? The answer is both. Whenever anyone is in cyberspace, she is also here, in real space. Whenever one is subject to the norms of a cyberspace community, one is also living within a community in real space. You are always in both places if you are there and the norms of both places apply. The problem for law is to work out how the norms of the two communities are to apply given that the subject to whom they apply may be in both places at once*.

A comparative review of legislation in the EU, the United States, Australia and New Zealand suggests that a principles-based approach is likely to offer the key to international interoperability of regulation (Toy 2013). It is also considered to be more flexible and therefore better able to accommodate future changes in technology. Lozada et al (2013, p.60) describe the technology used by marketers to gather personal data and the push by European and US authorities to protect consumer online privacy. The authors conclude:

> *Presently, there are no uniform international laws to deal with violations to online privacy. Thus, we are confronted with the dilemma of having no one specifically monitoring for potential violations to consumers' rights to privacy online. Last, it is becoming critical to ascertain the effectiveness of opt out options and their impact on alleviating consumer skepticism.*

At the time of writing the EU's General Data Protection Regulation 2012 had been passed by the European Parliament and was in the process of revision prior to publication of a final version. A number of commentators have examined the draft Regulation to consider how it might develop taking into account differences in approach between the EU and the United States, for instance (Schwartz 2013; Svantesson 2014).

SELF-REGULATION

Self-regulation by users of their use of the internet falls within the scope of user behaviour, which is considered later in this chapter (LaRose et al. 2003). Haufler (2001) has a useful definition:

> *Self-regulation occurs when those regulated design and enforce the rules themselves.*

One of the most prominent self-regulatory regimes affecting the social media services used in the UK is the U.S.-EU Safe Harbor arrangements with the Federal Trade Commission (FTC). Although this is based on an agreement between the European Commission and the US authorities, it is a voluntary arrangement with no external verification of registrations required. This has been reviewed previously and many concerns were raised about inconsistent and inaccurate registrations, as well as the very loose compliance requirements (Connolly 2008).

There has been some reaction against self-regulation following the financial crisis of 2008 and its aftermath, which was widely perceived as a failure of self-regulation (Slattery & Nellis 2011). Even before the crisis some commentators were suggesting that self-regulation is ineffective on its own (Collins 2006).

Cannataci and Bonnici (2003) examine self-regulation by Internet Service Providers (ISPs) and point out that self-regulatory approaches are constrained to some degree by national boundaries.

In the United States the role and actions of the FTC have been reviewed and it was found that the FTC has in effect codified certain norms and privacy best practices but that there has been a failure of self-regulation that has resulted in settlements with Facebook, Google and Twitter among others (Solove & Hartzog 2014; Hans 2012). The FTC operates through a number of divisions including the Division of Privacy and Identity Protection, a part of the Bureau of Consumer Protection. Under the Federal Trade Commission Act 2006, the FTC regulates unfair practices and has used this to make cases against SNS providers for misleading privacy policies, and gathering and distribution of personal data to other suppliers.

The European Union has seen a move to co-regulation, where the responsibility for regulation is placed on the regulated industry, but with supervision by the state. This is sometimes referred to as 'regulated self-regulation' and can be characterised as a combination of law-base or state regulation and measures taken by industry (self-regulation) (Hans-Bredow Institut 2006).

Privacy policies are another manifestation of self-regulation by industry. An initial investigation of the privacy policies of the largest SNS providers is described in Chapter 8.

CODE

Reidenberg (1998) first put forward the idea of 'Lex informatica' where the structure and architecture of cyberspace became a means of regulating it. In Code 2.0 Lessig (2006, p.5) states:

> *In real space, we recognize how laws regulate – through constitutions, statutes, and other legal codes. In cyberspace we must understand how a different "code" regulates – how the software and hardware (i.e., the "code" of cyberspace) that make cyberspace what it is also regulate cyberspace as it is.*

In other words: *"Code is law"*, in Lessig's catch-phrase. He goes on to argue that it is possible to build systems to reflect the values that are considered important or fundamental. Indeed it is impossible to avoid reflecting some values in the build, architecture or coding of cyberspace. Later he says: "*There is regulation of behaviour on the internet and in cyberspace, but that regulation is imposed primarily through code*." For example, "*code that encrypts regulates to protect privacy*". The code itself can be regulated (e.g. the US government requirement that it should have access to encryption codes built into any system) as well as being an agent of regulation. Lessig terms these East Coast code (e.g. laws enacted in Washington DC) and West Coast code (system design and architecture based in Silicon Valley).

Lessig goes on to look at the way in which code regulates privacy on the internet. Zittrain (2008) talks about loss of control over personal data. Lessig argues that the interests threatened by breach of privacy are diffuse and disorganised (unlike the copyright interests that are also affected by the internet). There are compelling interests arrayed against privacy interests, namely national security. Zittrain argues that this is reason why privacy legislation in the US is relatively neglected compared to copyright legislation. One solution might be Lessig's (2006, p.129) view that "*the protection of privacy would be stronger if people conceived of the right as a property right.*" This would mean that individuals would have control over the distribution and use of their personal information and they would attract payment for its use. A royalty system might need to be established to achieve this and the current economic model of free consumer use of social media may have to be reconsidered.

Recognition of the importance of system architecture in regulation is seen in the Information Commissioner's Office's (2008) 'Privacy by Design' initiative, an example of regulation by code.

The principle is that any system or service that is established should consider privacy issues from the initial design stage onwards, rather than being tacked on as an afterthought.

One way that 'code' is expressed is in the default privacy settings and the options that users are given for adjusting the level of disclosure that they are comfortable with. Willis (2014) questions the effectiveness of default positions (Track-Me / Do-Not-Track) as a way of regulating privacy on the internet. If rules (regulations) are imposed to mandate a Do-Not-Track default, for instance *"a firm's three main avenues for ensuring the failure of a default* [are]*: (1) utilizing transaction barriers, (2) harnessing judgment and decision biases, and (3) influencing preference formation effects"* (Willis 2014, p.112). She also considers it unlikely that competition between firms will lead to greater privacy. The implication of her work is that greater user awareness and education is the way forward.

NORMS

User behaviour is expressed collectively in societal norms and in market behaviour, as well as individually. Norms regulate services by creating default ways of working. For instance, there is a developing norm that users should be able to opt in to sharing their personal details with advertisers, rather than having to opt out. Instead of regulation Tene and Polonetsky (2012, p.66) argue that "*Policymakers should engage with this normative question, consider which activities are socially acceptable, and spell out the default norms accordingly*". They go on to suggest that risk analysis is a more rational basis for data protection than data minimisation (i.e. only collecting that data that is directly necessary for the original purpose it was gathered for).

Solove (2007, p.179) describes privacy in terms of social network theory and the norms that apply:

> *Social network theory often focuses primarily on connections, but networks involve more than nodes and links. There are norms about information sharing that are held within certain groups, such as norms of confidentiality*.

Regulation by the market (e.g. demand by consumers) is affected by privacy concerns as has been seen in the responses over changes to the privacy settings on Facebook (Helft & Wortham 2010). However some of this change may be due to pressure from regulators such as the Canadian Privacy Commissioner (Denham 2009).

Lessig (2006) dismisses the idea of the market as a means of protecting personal privacy on the internet. Other commentators have suggested that a market for personal data is beginning to develop (Yassine et al. 2012). This could in future become an effective means of regulating access to personal data, or a means of protecting personal privacy on social networks. Boyd (2012) asks about the degree of control users have to develop their own social norms. She identifies the use of real names as a norm on Facebook, which she attributes to Facebook's origins as a campus-based service. When the service expanded beyond the campus not all users adopted this norm. When Google+ attempted to force users to register with their real names, there was a consumer backlash, which suggests that there may be a different norm for Google+ users (Boyd 2012).

Is some form of feedback mechanism possible so that systems regulate themselves? Parallels can be seen in biological regulation where for instance increased blood sugar leads to a rise in insulin production which lowers the blood sugar level, which in turn is detected by the body which responds by reducing the production of insulin. An auto-regulatory system (not to be confused with self-regulation) might be an interesting avenue to explore and may be one of the characteristics of regulation by the market (Lessig 2006).

The Information Commissioner's Office (2014) in the UK has actively educated users to take responsibility for protecting their own privacy. User education was also seen as one strand of an effective strategy for avoiding some of the risks associated with online SNSs. This is a reflection of Dry's (2007) historical analysis of risk and regulation and the benefits of individuals taking greater responsibility for risk assessment. This could be seen as a possible mechanism for regulation by the market. Users' behaviour is modified in response to perceived risk, affecting the demand for SNSs. Nosko et al (2012) found that user disclosure behaviour changed in response to exposure to stories about data misuse and this might be an avenue for future user training.

Social norms can act against prudent behaviour as shown in a study where a fictional Facebook account was established. The research showed that people were willing to 'friend' the fictional account and speculated that this might be due to norms of reciprocity (Lemieux 2012). However Hooper and Kalidas (2012) found that there were clear views about unacceptable behaviour among a sample of youths (18-20 years old) on Facebook. Their evidence suggests that acceptable behaviour (about which there is less clarity) may be subject to herding, so that what is considered acceptable is more dependent on the context. Martin (2012) tries to

quantify social norms by examining user responses to a series of vignettes and concluded that norms depend on context.

Nunan and Yenicioglu (2013) in their examination of social media in market research (and advertising) highlighted the problem of lack of informed user consent and the ethical problems that this presents. They point out that regulation lags too far behind technology changes and that there is a danger of alienating users. They suggest that users need to be more involved in the research process and that protecting them from harm is a major consideration.

MIXED APPROACHES TO REGULATION

Lessig (2006, p.267) argues that combinations of the four modalities protects the individual: "*Here again, then, the solution is a mixed modality strategy. As LAW creates the incentive for a certain change in the CODE of spam (it now comes labelled). That law is enforced through a complex set of MARKET and NORM-based incentives*".

Solove (2007, pp.190, 191) draws parallels between privacy and copyright and suggests that there has to be a balance between freedom and control of information. He concludes that "*We are witnessing a clash between privacy and free speech*" and suggests that (in the United States) "*we can rework the law to make it a useful instrument in balancing privacy and free speech*". He suggests that lawsuits provide a middle ground between a libertarian approach (leaving law out of considerations of privacy) and "*An authoritarian approach which involved direct restrictions on internet expression*". In his view the law should recognise the variety of situations where privacy is a consideration and it "*should also increase its recognition of duties of confidentiality*". Social norms are a powerful means of regulating behaviour and can be a strong alternative to resorting to legal remedies. To some extent social norms also drive legislation by placing pressure on law-makers as well as on those administering the law. Finally Solove advocates parties working out disputes between themselves without resorting to the law and refers to Lessig and Reidenberg's view that technical architecture can be used to protect privacy (Solove 2007; Weber 2002).

Spinello (2006) takes up Lessig's theme of modes of regulation and questions whether special regulation is needed for the internet, the so called 'law of the horse' (i.e. if special laws are not required for horses, why is one needed for the internet?). Spinello concludes that special legislation is not required in this case.

Schmidt and Cohen (2013, p.66) in their forward view of humans and technology talk about different modes of regulation (*"coping strategies: corporate, legal, societal and personal"*).

Although they do not map neatly onto Lessig's regulatory modes, there is a correlation between their categories and the model used in this research.

So what of the United Kingdom? Buckley Owen et al (2012) point out that the lack of a national information policy in the UK means that responsibility for regulating information and communication services falls across several different agencies. The one most directly concerned with regulating access to personal data on social networks is the Information Commissioner's Office, under the terms of the Data Protection Act 1998 and discussed in detail in Chapter 7.

Lofstedt et al (2011) argue that the relationship between public and regulators has changed over the last 20 years. Several major scandals have led to greater public distrust. This has resulted in a shift from a consensual style of regulation to a model based on public participation, transparency, and more powerful non-governmental organisations (NGOs).

Lessig's (2006, p.232) view is that a combination of law, norms and architecture/code is the most effective means of regulation of privacy on the internet. He does not believe the market is an effective regulator in this instance and concludes:

> *Collective action must be taken to bend the architectures towards this goal* [protecting privacy]*, and collective action is just what politics is for. Laissez-faire will not cut it.*

This view is open to challenge, especially in light of emerging views about a market for tradeable personal data. However it is difficult to evaluate these different approaches without some way of measuring their effects. The purpose of this research is to develop a methodology, or at least an approach based on risk analysis to compare these modes of regulation.

MEASURING REGULATORY EFFECTIVENESS

A starting point for a review of regulatory effectiveness might be to look at its effect on risk. Hutter (2005) documents the move towards risk-based regulation in the UK and Europe but points out that risk is often subjective and therefore difficult to define precisely. The author goes on to talk about the use of impact and probability estimates as a basis for prioritising risk.

This research uses the measurement of risk as a tool for evaluating regulatory effectiveness. This is based on the premise that one of the purposes of regulation is to manage risk. One of

the problems with this is that there are different views of what risk is. Aven and Renn (2009) discuss risk in terms of something of human value and suggest that both uncertain events (likelihood) and outcome (impact) need to be taken into account in assessing risk. Different methods of assessing regulatory impact were introduced to improve the reputation (and effectiveness) of EU regulation (Torriti 2007).

The idea of using risk measurement for assessing regulatory impact has been around for some time. This research is based on the premise that risk assessment provides the foundations for evaluating regulatory effectiveness. The British Computer Society described compliance with the Data Protection Act 1998 in terms of risk management (Room 2007). However not all commentators are convinced about the value of risk-based legislation. Haythornthwaite (2006) takes the view that where the risk is confined to an individual and does not harm wider society, it should become their individual responsibility, for instance by taking out insurance.

Heyvaert (2011) goes further in suggesting a shift in risk regulation (based on work on climate change) and a move to integration and orchestration and away from individualisation and compartmentalisation. Rather than looking at the risks associated with use of online SNSs, we may need to consider wider risks associated with internet use. For instance, the risks associated with gathering and exploitation of behavioural data or the inadvertent download of malware are systemic risks of internet usage. Risks associated with making personal data available to a range of providers (not just SNSs) leave the door open to abuse of that data.

Liu and Terzi (2010) have developed a 'privacy score' that "*measures the user's potential privacy risk due to his or her online information sharing behaviour*". Their mathematical models are intended to estimate the sensitivity and visibility of personal information. However the models apply to individuals' attitudes to sensitivity of different types of personal data, the nature of the network, and the size of individual personal networks online. It is difficult to see how different regulatory modes would affect the scores that are derived from this model.

Swedelow et al (2009) have constructed a universe of almost 3,000 risks over a 35-year period to create a means of evaluating European and US regulation. This provides a resource for theory testing and theory building and serves as a potential test-bed for analysing risks associated with use of online SNSs.

CONCLUSION

The literature survey focused on: risks associated with personal data on SNSs; ways in which access to personal data is regulated; and on assessment of regulatory effectiveness.

There is a well-developed literature on general risks and descriptions of specific risks faced by users of social media, but no systematic analysis of the ranges of risk associated with SNSs. Perceptions of internet risks are covered by the OXIS longitudinal surveys (Blank 2010; Dutton & Blank 2011; Dutton & Blank 2013). However they do not present a systematic classification of risk and there is a need for a typology of risks associated specifically with SNSs.

The literature has established that one of the purposes of regulation is to manage risk. Lessig's (2006) model of regulation is a starting point for the regulation of SNSs, although this predates the emergence and rapid expansion of social media. Wu (2010) provides a good historical analysis of regulation of digital and electronic communications services in the United States and Zittrain (2008) provides a forward view of regulation of the Internet. There is also considerable literature on statutory regulation particularly around data protection in Europe, although very little on its application to SNSs. Privacy by design (an aspect of Lessig's 'code') is also well covered. Norms have been considered by US-based researchers such as Boyd (2010) and UK researchers such as Miller (2011). However there has been relatively little on UK regulation since Cooke (2004) and Oppenheim (2001) both of whose published work in this area predates SNSs. The Lessig (2006) model deserves further scrutiny to assess its suitability for the current social media landscape and applicability to SNSs.

The literature on regulatory effectiveness tends to consider its effects on the markets or on businesses. The link between risk and regulation is well established, but there is no systematic research on the use of risk measurement to evaluate regulatory effectiveness. This research aims to fill the gap by: identifying and categorising the risks associated with social media; refining the model of regulation that can be applied to SNSs; and developing a method in which risk concepts can be used to compare and evaluate the relative effectiveness of different regulatory modes.

CHAPTER 3 – METHODOLOGY

METHODS IN INFORMATION SCIENCE RESEARCH

"The primary goal of research is to link the empirical and the theoretical". (Ragin 1994)

This review of regulation calls on a variety of methods from sociology, economics and legal studies. The emphasis is on qualitative methods to identify issues and argumentation supported by quantitative studies of behaviour and attitudes. Textual analysis is particularly relevant in exploring privacy policies of online SNS providers. This research is intended to influence information policy relating to personal data and is informed by the range of research methods used in information science.

Hjørland (2002) uses domain analysis to define the scope of information science. Specifically he identifies eleven domains of which two are of direct relevance to this study: *"empirical user studies"*; and *"studies of structures and institutions in scientific communication"*. However he concludes that:

> *… empirical studies of users may represent an important approach to domain analysis in IS if they are informed by proper theory. They may, for example, provide information about differences in information needs in different communities. They should be combined with other approaches, including:*
>
> - *bibliometric studies;*
> - *epistemological and critical studies; and*
> - *studies of structures and institutions in scientific communication.*

(Hjørland 2002, p.432)

The last of these headings makes uncomfortable reading because of the emphasis on "*scientific communication*". He bases this on the UNISIST model of the main categories of information sources and the key players in this. The model breaks down when applied to online SNSs, where one of the characteristics of these services is user-generated content. The agencies that are responsible for managing and distributing this information do not fit easily into the categories of: publishers; abstracting and indexing services; libraries; Information centres; and clearing houses, although data centres could apply. Therefore a new model is needed if this research is to be considered a part of the domain of information science.

Robinson (2009) in setting out "*to derive a conceptual model for information science, which is both academically sound and practically useful*" extends Hjørland's model of information science to a wider-ranging definition: "*Information science can be understood as a field of study, with human recorded information as its concern, focusing on the components of the information chain, studied through the perspective of domain analysis, and in specific or general contexts*." Personal data on online SNSs can be taken to be human recorded information (including behavioural data from tracking cookies and web beacons). Components in the information chain include SNS stakeholders: users; service providers; and regulators. Although the primary focus of this work is on risks to users, the roles of the other actors (or components) in the information chain are also considered.

Rowlands, Eisenschitz and Bawden (2002) consider the limitations of the political economy frame for the study of information policy and suggest that it is inadequate in light of non-market conceptions of information and the emergence of human rights legislation. They suggest that the study of information policy fragments into: laws, regulation, IM practices, and institutional cultures. They suggest that "*inquiry into the political realm can never be value-free*" and that there are no objective truths in information policy. The frame used to analyse information policy is a set of values and concepts people use to make sense of the world around them.

The research sets out to test the hypothesis that risk can be used as a means of evaluating the relative effectiveness of different modes of regulating access to personal data on SNSs. This in turn will help to throw light on question of whether legislation alone is the best way of regulating access to personal data on online SNSs accessed by users in the United Kingdom.

Approach

An initial view of user risk would suggest that an empirical approach to measuring risk should form the basis for comparison of different regulatory modes. However many of the risks identified are difficult to measure objectively, despite attempts in this research to do so. The research may be limited to a qualitative evaluation of perceptions of risk or a quantitatively indicative approach. So, for example, regulatory effects can be described in terms of their tendency to increase or decrease the probability of a risk occurring and their tendency to increase or decrease the impact of the risk event if it occurs.

Another avenue to explore would be the effect of regulation on perceptions of risk by users and the response of different stakeholders to regulation or potential regulation. These

approaches call on well-established social science research methods including ethnography, case studies and attitudinal surveys. Outhwaite and Turner (2007) suggest that social science methodology has two meanings:

> *methodological issues arising from and related to theoretical perspectives, as in Marxist, functionalist or feminist methodology*

and

> *issues of specific research techniques, concepts and methods*.

This research focuses on the second of the two definitions.

McNeill (2005) maps the development of social science research from thinkers such as Marx, Durkheim and Weber as well as the early social research of Charles Booth in London in the late 19th and early 20th Centuries (Charles Booth Online Archive 2013). He goes on to describe the Chicago School which pioneered anthropological research, through to the research by feminist scholars in the 1980s.

The methods adopted for this policy research were largely qualitative. A particular challenge of social research is combining the insight provided by qualitative research with the generalizability of results that comes from statistically significant quantitative work. Much of the discourse on research methods attempts to reconcile these two approaches (Niglas 2010; Turrow 2004).

In his chapter on bridging the quantitative and qualitative divide Turrow (2004) describes process tracing that uses qualitative analysis focused on processes of change within cases to uncover causal relationships. Underlying quantitative findings focus on tipping points – explaining points in time-series data where changes occur. He suggests that sequencing qualitative and quantitative approaches in a single study allows triangulation of results and provides better insights into the phenomenon being investigated.

Both qualitative and quantitative approaches were used for this research. The qualitative results (e.g. from semi-structured interviews and examination of legislation) are based on textual analysis to identify themes and topics which were then further explored. The data provided material for development of a working hypotheses about the relative effectiveness of different types of regulation.

Whilst it is important to acknowledge any assumptions or starting points for the research, an approach based on Glaser and Strauss's Grounded Theory seems more appealing, where an initial hypothesis is generated from a pilot study and this is further investigated using empirical data to test the hypothesis. This method is described in detail by Charmaz (2006) who states that "*theoretical sampling is less of an explicit procedure than a strategy that you invoke and fit to your specific study*".

This thesis also examines other empirical data such as records of court cases to test the working hypotheses. This works on the principle of falsifiability of an hypothesis (Popper 1959). In other words, the research questions are framed in such a way that it is possible to disprove the hypothesis with empirical data if it turned out to be false.

Ragin (2000) combines qualitative and quantitative techniques in 'fuzzy sets' where individual entities have varying degrees of membership of a set. So for example, one could study varying degrees of risk associated with personal data or applying to users. He goes on to suggest that if Theory 1 proposes cause and effect and Theory 2 proposes multiple causes of the same effect, Theory 1 may still be valid. Applying Ragin's principle to this research one could posit that data protection legislation will result in lowered risk to users of SNSs. This hypothesis is not necessarily undermined by the view that several different modes of regulation working in concert may be even more effective in lowering the risk to users.

METHODS USED

The research methods used can be broken down into a series of distinct areas of activity designed to throw light on specific aspects of risk and regulation in SNSs:

- Survey of attitudes to risk
- Review of legislation
- Investigation of privacy policies
- Interviews with stakeholders
- Identification and measurement of risk
- Case studies of SNSs
- Assessment of regulatory modes

SURVEY OF ATTITUDES TO RISK

A preliminary qualitative survey of individual users (Appendix B) and professionals responsible for data protection identified the following issues:

1. Mixed views about the efficacy of data protection legislation as a means of regulating access to personal data
2. Concern about lack of enforcement or difficulty of enforcement of data protection legislation across national boundaries
3. The need for SNS providers to be more open about what they do with personal data, and defaulting to more secure settings
4. The need for greater education and awareness of the risks associated with posting personal data on social networks
5. View that better encryption standards will help to protect personal data

These themes were followed up in an online survey of LIS professionals in the UK (Appendix C). This particular group was selected for the following reasons: the investigation is focused on regulation that applies to UK users of SNSs; LIS professionals' knowledge and use of SNSs as intermediaries, to communicate with their customers, and to educate; and knowledge of this group by the researcher. The survey used convenience sampling and therefore cannot be extrapolated to represent the entire group. However the 222 responses allowed a quantitative analysis of the survey results to gain insight into perceptions of risk and priorities. SurveyGizmo was used to deliver the survey and gather the results for analysis. The data analysis was primarily conducted using Microsoft Excel spreadsheet software for the numerical data and QSR NVivo 10 for content analysis of the responses to the open questions.

REVIEW OF LEGISLATION

A key component of this research was a systematic study (Chapter 7) of the legislation that applies in the UK to protect users against the risks associated with putting personal data up on SNSs. This examined: the Data Protection Act 1998; the Data Protection Directive (95/46/EC); the U.S.-EU Safe Harbor framework and the EU General Data Protection Regulation 2012. The research also identified relevant secondary legislation including statutory instruments and European Commission opinions and statements.

This investigation is based on 'black letter law' research methodology. This involves reviewing legal sources and related material to address the following questions:

1. What are the rules governing access to personal data on social networks
2. How did the rules come about?
3. What is their effect?
4. What can be changed?

Another avenue explored was court cases related to release of personal data on online SNSs. However there was not a sufficiently large body of cases to estimate economic costs in the UK or even across the whole EU. However the small number of cases that do exist provided additional insights into the risks and issues that arise.

The Information Commissioner's Office was approached for guidance on work that they may have commissioned in this area which has not already been published. Existing and proposed EU legislation was analysed and regulators and legal experts were interviewed to ascertain whether there had been any relevant work done on quantification of risk.

INVESTIGATION OF PRIVACY POLICIES

The privacy policies of the largest, English-language SNSs available to UK users were reviewed (Chapter 8). The policies of eleven major service providers were included. Each of the selected SNSs was thought to have more than 10 million registered users globally (Wikipedia 2014).

QSR's NVivo10, a content analysis tool, was used to mark up, code and analyse the text of the privacy policies and to identify common themes. Rather than starting with a pre-set coding frame, the coding was developed from an analysis of the text, a method known as 'literary warrant'. Once the coding frame had been developed this was applied to all the policies. As new themes emerged they were back-indexed. The research questions were also revisited to ensure that the coding reflected the scope of this research.

INTERVIEWS WITH STAKEHOLDERS

Data protection experts including regulators, representatives of the online advertising industry and privacy campaigners were interviewed (Appendix D). This strand of the research provided an opportunity to explore specific aspects of regulation.

A series of semi-structured interviews was undertaken, with the aim of eliciting responses from a diverse group of respondents. The following topics were explored:

- Attitudes to risk associated with social networks, usage of social media and view on effectiveness of different types of regulation of access to personal data
- Users of personal data (e.g. advertisers) and the extent to which their behaviour is affected by UK legislation – in comparison with other modes of regulation
- Views of social network service providers on effectiveness of legislation in the UK and other forms of regulation

- View of regulators and what they perceive to be the challenges for future regulation of access to personal data

Semi-structured interviews were recorded, transcribed and sent to the interviewees for checking. In all but one case permission was given to provide attributable quotes. A set of open questions was devised to tease out the issues surrounding the development of data protection legislation, shortcomings and ways in which it might develop in future. Active note-taking was also used to supplement the audio recordings. The content of the interviews was coded and analysed using NVivo10.

IDENTIFICATION AND MEASUREMENT OF RISK

The first part of this research was to identify risks starting from the universe of general risks compiled by Swedlow and associates (2009). The research also considered specific risks associated with SNS use and reported in the literature. This was enriched with additional analysis of risks identified during the main online survey and developed further during interviews with regulators and other stakeholder representatives (Appendix D). These fed into the development a risk model based on a typology of risk developed specifically from the perspective of individual users of SNSs.

Two dimensions of risk were considered when attempting to measure or quantify risk:

1. the **probability** of occurrence; and
2. the **impact** that risk has if it comes to fruition.

Some methodologies combine these into a single risk value – this is often possible where the risk impact can be measured in terms of numbers of occurrences, or numerical values such as economic cost or benefit. One of the challenges of this stage of the research was to determine whether it was possible to assign a numerical value to the impact of allowing access to personal data on online SNSs.

SNS VIGNETTES

The final stage of the data gathering involved investigation of three online SNS services to consider the ways in which 'privacy by design' is implemented and the choices offered to users. These were intended to provide an in-depth view of the factors that underpin the privacy policies of SNS providers. The SNS vignettes used case study methods and were based on participant observation, where the researcher's own SNS profiles were used to identify user choices and their effects on visibility of personal data. The results of this

investigation are discussed in Chapter 10. Case study methods call on hermeneutics, which involves a deep understanding (known as '*verstehen*') of the context of the situation or phenomenon being studied (Brady & Collier 2004).

Ragin and Becker (1992) identify four types of case study:

> *1. Cases are found – empirically real and bounded, but specific. Identified and established as cases in the course of the research process*
>
> *2. Cases are objects – Cases are real and bounded but general and conventionalized. Cases defined by the literature – e.g. nation-states*
>
> *3. Cases are made – "theoretical constructs that coalesce in the course of the research" - see emerging patterns and construct theories around them*
>
> *4. Cases are conventions – "general theoretical constructs,* [...] *as products of collective scholarly work and interaction...external to any particular research effort*

The investigations of privacy policies of SNSs and of 'privacy by design' fall into the first category of case study – "*empirically real and bounded but specific*". The research used participant observation (ethnography), which is seen as an appropriate methodology for studying "*complex social relations and organisational processes*" (Delbridge & Kirkpatrick 1994).

Miller (2011) uses local ethnography and comparative anthropology for his study of Facebook based on case studies conducted in Trinidad. His starting point was that each experience of Facebook is unique. He sets this in the cultural context "*but what any given population actually uses, based on that facility, quickly develops its own local cultural genre and expectations which will differ from others*". He goes on to say that "*there is no such thing as Facebook from the perspective of cultural relativity. Facebook is only the aggregate of its regional and particular usage*". He posits fifteen theses on what Facebook might be (based on user experiences) and this in turn may affect attitudes to personal data and who should have access to it under what conditions. These are themes that were explored in this study.

A protocol for the online case studies was developed in order to systematically investigate the privacy settings of three SNSs (Facebook, LinkedIn and Twitter. Although these vignettes are not full case studies, this approach draws on case study methods (Appendix E). The online

sessions were recorded by means of screen capture (using the Windows Snipping Tool) and by making notes as a commentary about the actions taken. In the case of Facebook, NCapture for NVivo was used with the MS Internet Explorer software to record the metadata associated with postings on a Facebook site.

ASSESSMENT OF REGULATORY MODES

The effectiveness of the four regulatory modes for reducing risks was assessed by looking at how regulation affects users. The four regulatory modes are based on Lessig's (2006) model of regulation of the Internet, and adapted to encompass the range of regulatory approaches used for SNSs (discussed more fully in Chapter 6):

1. Legislation
2. Self-regulation
3. Code
4. Norms

The potential parameters for comparison of the regulatory modes are:

- Adaptability to changing technology
- Comprehensiveness
- Which risks addressed
- Effects on risk
- Impact on stakeholders
- Cost of enforcement
- Benefits (savings) from enforcement (i.e. consequence of not regulating)
- Alternatives available

Legislation was used as the point of reference for comparison with other regulatory modes. This refers back to the original question: Is UK law-based regulation the most effective way of protecting users against the risks associated with use of SNSs? In the UK, the Data Protection Act 1998 (DPA) is the main instrument for legislative regulation of access to personal data and is the most visible and widely-applied regulatory approach to protection of personal data in the UK.

A detailed textual analysis of existing primary and secondary legislation in the UK informed the choice of respondents and questions for discussion with the regulators and legislators. The interviews explored the issues that specifically relate to SNSs. Interviews with legislators and

regulators considered possible future intentions and directions for regulation. The question areas were:

> What measures are in place to deal with social network providers/services?
> What are the required measures to effectively regulate this area?
> What approaches might be adopted by the regulators in future?

This is a form of socio-legal research, distinct from the scientific research method. Feldman (1989) in 'The Nature of Legal Scholarship' states that "*scholarship is related to the good of knowledge*". He suggests that the scientific method is not appropriate for legal scholarship for the following reasons:

- The limitations of the scientific method – observation may change the thing being observed, difficulty in constructing a null hypothesis when there are many possible hypotheses to test. The inappropriateness of the falsifiability rather than verifiability of hypotheses because of the impossibility of constructing a null hypothesis for most social science questions.
- His second objection is that "*claims to scientific objectivity are often inflated*". By this he means that study of law and legal systems is not value free and that objectivity is therefore difficult to attain in legal studies.
- Thirdly, he states "*In refusing to use legal techniques, either to investigate that claim or to discover the state of the law, one discards analytical tools of some interpretative value*." In other words he is suggesting that legal research methods may be more appropriate than scientific methodologies for research in this domain.

Feldman (1989, p.503) goes on to suggest that scholarship can be evaluated "*as being more or less in tune with certain formal values which are integral to a serious search for truth. These include*:

> *(1) a commitment to employing methods of investigation and analysis best suited to satisfying that curiosity; (2) self-conscious and reflective open-mindedness...; and (3) the desire to publish the work for the illumination of students, fellow scholars or the general public and to enable others to evaluate and criticise it."*

Research Ethics and Project Risks

The main ethical issues are associated with surveying individual attitudes and perceptions. Data gathered was kept securely and the results were anonymised to prevent the possibility of identification of an individual based on their reported profile. Data that identifies individuals has been removed from the questionnaire survey results. Interviews with legislators and regulators are difficult to anonymise, so respondents were given the opportunity to review the interview notes and any quotes proposed before they were analysed.

There were a number of project risks associated with this research that had to be taken into account (Table 2):

Table 2 - Project risk management

Risk	Response	Result
Lack of access to the right people – especially service providers and legislators, undermining the validity of the findings	Minimise probability: Early identification of targets. Establish contact at conferences and meetings	Get into diaries of busy and difficult to access people Establish alternative, authoritative contact points
Difficulty in constructing a representative sample of stakeholders for survey, making it difficult to draw conclusions about attitudes in the general population	Avoid: construct a qualitative survey that obtains a wide range of views	Provides a qualitative response, which may not be regarded as well as a quantitative survey
Difficulty reconciling different models of risk, makes it difficult to establish a testable proposition for this research	Minimise probability: Evaluation of different models before picking one or developing a new one. Concentrate on testing the selected model	New model of risk developed. Greater focus for the research by having a single testable model to work with
Difficulty finding a means of quantifying risk making it difficult to test the relative effectiveness of different modes of regulation	Minimise impact: use qualitative measures of effect on risk to evaluate different regulatory modes	Development of a new method of assessing regulatory effectiveness

Risk	Response	Result
Inconclusive results make it difficult to make positive assertions	Minimise probability by framing questions in such a way that it is possible to make a positive assertion whether the results are positive or negative	Conclusions expressed in such a way that it is possible to publish the results

CONCLUSION

This research calls on a variety of methods from sociology, economics and legal studies. It falls within the domain of information science research, which has adopted many of the methods used in sociology and ethnography. This study has mostly used sociological research methods based on an inductive approach. This is informed by Weber's response to the limitations of logical positivism (Outhwaite & Turner 2007). Instead he bases his new approach on human cultural values and norms. Popper (1959) and Kuhn (1970) also proposed a post-positivist approach which codifies observations. Throughout this research a data-driven approach was adopted. Examination of data from the preliminary phase of the research allowed construction of hypotheses using an approach based on Grounded Theory (Glaser & Strauss 1967). The main phase of the research gathered new sets of data (an empirical approach) to test these hypotheses.

The context of social networking services is important to get a fuller understanding of regulation and allows an 'interpretivist' approach as suggested by Weber and others (Gerth & Mills 1970; Outhwaite & Turner 2007). This was particularly relevant for the semi-structured interviews with industry experts and the qualitative survey of LIS professionals to identify risks and potential responses. These research methods were used to see what emerged. The findings were analysed to identify patterns and to construct models that could be used in the evaluation of different regulatory modes, a classic inductivist approach.

The socio-legal aspects of this research also focus on: formal values; the subjective nature of law; and the use of the interpretivist tools that have been developed in legal research.

Figure 4 shows the relationship between the research methods adopted for this study. During the initial phase, a preliminary survey yielded several lines of investigation. The three research lines were supported by the literature survey and provided data for the formulation

of the hypotheses, and the construction of some outline conceptual models for use in the main research phase.

Eventually five different research lines were established for the main phase of research, three being carried forward from the preliminary phase. This 'triangulation' allowed a deeper understanding (or 'verstehen') of regulation and SNSs.

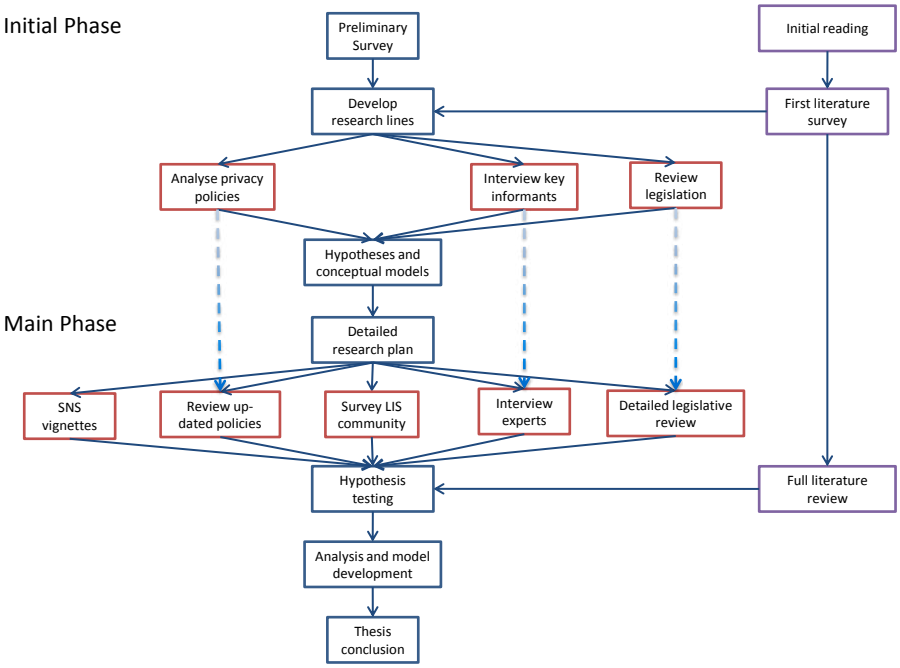

FIGURE 4 - RESEARCH METHODS USED

SECTION II. RISK

CHAPTER 4 – PERSONAL DATA, PRIVACY AND DATA PROTECTION

INTRODUCTION

This chapter sets out to identify personal data in the context of SNSs and to explain why it is different to other classes of data. It considers different types of personal data and how it is used in SNSs. This chapter also discusses the concept of data protection and draws distinctions between it and information privacy.

WHAT IS PERSONAL DATA?

Article 2(a) of the EU Data Protection Directive (95/46/EC) defines personal data in the following terms:

> *(a) 'personal data' shall mean any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;*

Section 1(1) of the Data Protection Act 1998 uses a more useful definition in that it is driven by purpose rather than trying to enumerate the different types of personal data that exist:

> *personal data" means data which relate to a living individual who can be identified—*
>
> *(a) from those data, or*
>
> *(b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller, and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual*

Milliard and Hon (2012) discuss the potential problems that arise from the discrepancies between the UK's and the EC's definitions of 'personal data'. There is a concern among other EU states that the UK's definition of personal data is too narrow, whereas the UK concern is that the EU definition is too prescriptive and insufficiently flexible to deal adequately with new cases or technologies that arise. For the purposes of this study it will be necessary to take both definitions into account, because they both affect the practices of the SNSs that are used by UK residents.

Schneier (2010) suggests that different types of personal data on social networks (described in Chapter 2) will require different measures to protect them. Section 2 of the Data Protection Act does identify some data as being sensitive and therefore warranting greater protection:

> *In this Act "sensitive personal data" means personal data consisting of information as to—*
>
> *(a) the racial or ethnic origin of the data subject,*
>
> *(b) his political opinions,*
>
> *(c) his religious beliefs or other beliefs of a similar nature,*
>
> *(d) whether he is a member of a trade union (within the meaning of the Trade Union and Labour Relations (Consolidation) Act 1992),*
>
> *(e) his physical or mental health or condition,*
>
> *(f) his sexual life,*
>
> *(g) the commission or alleged commission by him of any offence, or*
>
> *(h) any proceedings for any offence committed or alleged to have been committed by him, the disposal of such proceedings or the sentence of any court in such proceedings.*

BEHAVIOURAL DATA

A great deal of recent attention in the press has been devoted to behavioural data, which users may not have provided explicit, informed consent for SNS providers to use. The permission may be hidden in the terms of service that a user signs up to in order to gain access to the service. This information may then form the basis for intrusive advertising directed at the user when they open their browser. A series of articles in the Wall Street Journal by Julia Angwin and colleagues describes tracking technologies and some of the issues that arise, including privacy concerns (Angwin & Mcginty 2010). There is a view in the advertising industry that this is not personal data, because an individual cannot be identified – particularly for aggregated data. However other researchers have demonstrated that it is possible to de-anonymise data to identify individuals (Narayanan & Shmatikov 2009; Federal Trade Commission 2010). Therefore the view has been taken that behavioural data is personal data and is included in the scope of this study.

IP ADDRESSES

Defining personal data may appear to be straightforward initially, but has resulted in discussion about what is personally identifiable information (Millard & Hon 2012; Gilbert 2014; Krishnamurthy & Wills 2009; Boyd 2012). For instance, an IP address is regarded in some quarters as personal information, because it is data that is often associated with an individual user and can in some instances be used as a proxy identifier. The Advertising Standards Authority suggests that online behavioural advertising (OBA) does not use personal data and that therefore some of the complaints it receives are outside its remit (ASA 2013):

> *Some complainants believed that OBA used personal data and that OBA infringed their rights to privacy and basic human rights.*

However there has been ample demonstration of the fact that cookies can be combined with other publicly available data to reveal information about individual online activity (Information Commissioner's Office 2012; Narayanan & Shmatikov 2009). For these reasons cookies and IP addresses are included in the scope of this chapter.

THE IDEA OF PRIVACY

Central to the discussion about regulating personal data is a question about what is privacy. Habermas and Burger (1989, pp.43–46) suggest that the distinction between public and private spheres arose from the development of an affluent, educated urban class from the mid-eighteenth century onwards. Habermas and Burger (1989, p.159) go on to talk about the "*Mutual infiltration of public and private spheres*" and observe:

> *The shrinking of the private sphere into the inner areas of a conjugal family largely relieved of function and weakened in authority – the quiet bliss of homeyness* [sic] *– provided only the illusion of a perfectly private personal sphere; for to the extent that private people withdrew from their socially controlled roles as property owners into the purely "personal" ones of their noncommittal use of leisure time, they came directly under the influence of semipublic authorities, without the protection of an institutionally protected domestic domain.*

In defining privacy Warren and Brandeis (1890) quoted from Judge Cooley, who spoke about privacy as: "*the right to be let alone*" in response to the growth of photo-journalism in the United States. Their argument was that common law provides a basis for protection of personal privacy. They stated "*The common law secures to each individual the right of*

determining, ordinarily, to what extent his thoughts, sentiments, and emotions shall be communicated to others". They went on to state that "...*the individual is entitled to decide whether that which is his should be given to the public*".

In Warren and Brandeis's (1890) view there are limitations to the right to privacy, particularly in relation to public good and they acknowledge that "*To determine in advance of experience the exact line at which the dignity and convenience of the individual must yield to the demands of the public welfare or of private justice would be a difficult task*". They argue that many of the laws already then available afford degrees of protection against invasion of privacy – such as copyright law for publication of photographs or literary works, or the tort of breach of trust by publishing confidential information, or breach of implied contract where private information made available in the course of delivering a service is then made public. They enumerate some useful principles that help to define the scope of what they mean by 'the right to privacy':

> *The right to privacy does not prohibit any publication of the matter which is of public or general interest.*
>
> *The right to privacy does not prohibit the communication of any matter, though in its nature private, when the publication is made under circumstances which would render it a privileged communication according to the law of slander and libel.*
>
> *The law would probably not grant any redress for the invasion of privacy by oral publication in the absence of special damage.*
>
> *The right to privacy ceases upon the publication of the facts by the individual, or with his consent.*
>
> *The truth of the matter published does not afford a defence.*
>
> *The absence of "malice" in the publisher does not afford a defence.*

One of the earliest privacy cases in England and Wales was *Prince Albert v Strange*, where the Prince Consort sought to prevent the publication of a catalogue of etchings made by him and the Queen. The High Court found:

> *The question here is, how far the publication of this Catalogue is in violation of the law? That there is property in the ideas which pass in a man's mind is consistent with all the authorities in English law. Incidental to that right is the right of deciding when and how they shall first be made known to the public. Privacy is a part, and an essential part, of this species of property. In Millar v. Taylor, the property which a man had in his unpublished ideas was admitted by all the Judges: Donaldson v. Beckett.*

Since the *Prince Albert v. Strange* case some commentators have distinguished between 'information privacy' and wider definitions of privacy. For example Westin (2003) includes in his definition of privacy: "*the claim of an individual to determine what information about himself or herself should be known to others*".

There have not been many cases in the UK courts about breaches of privacy or other damages resulting from disclosure of personal data on SNSs. In *Applause Store Production Ltd & Anor v Raphael*, where the defendant had created a personal profile of the claimant on Facebook containing some true information and some defamatory material the court found:

> *It is reasonably clear that damages in cases of misuse of private information are awarded to compensate the claimant for the hurt feelings and distress caused by the misuse of their information: see for instance McKennitt v Ash [2006] EMLR 178 [162].*

Malhotra et al (2004) looking at internet users' information privacy concerns developed a multi-dimensional grid of privacy, following surveys of internet users, based on:

- Collection of data – gathering of individual specific data by sites
- Control – whether users have the choice of opting out, for instance
- Awareness of privacy practices – knowing how the data will be used

An IP address, which can be attributed to an individual, can be regarded as personal data:

> *If one considers the definition of personal data provided in Article 2 of Directive 95/46/EC, 'any information relating to an identified or identifiable natural person (data subject); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number', it is only possible to conclude that IP addresses and the information about the*

activities linked to such addresses constitutes personal data in all cases relevant here.

(Hustinx 2010)

Westin (2003) suggests that privacy issues are different in democratic and authoritarian societies. It is not clear whether this is a qualitative difference or a matter of degree. It could be argued that in both types of society there is a balance between individual autonomy and public good. If the UK is assumed to be a broadly democratic society, then the balance would tend towards individual autonomy. However, the UK's Draft Communications Data Bill 2012 (Cm 8350) is perhaps a reflection of the move towards a more authoritarian society with greater emphasis placed on security at the expense of individual privacy. The Bill makes provision for monitoring of social networking activity as well as other telecommunications traffic that can be used to counter terrorist plots.

According to Lessig (2006) there are three different concepts of privacy that affect any considerations of regulation in this area:

- Preserving dignity
- Minimising intrusion
- Constraining the power of the state to regulate

Solove (2007) argues that information confined to a group should be considered private, even if that group is quite large. This could also apply to groups on the internet. In his view if a member of the group makes the information available beyond the group that is a breach of privacy: "*when information is contained within a particular group and a person causes it to leap the boundary, then this is a privacy violation even if the original group is large*."

An alternative view suggests that too much privacy is a bad thing because it inhibits delivery of free services to users. Hammock and Rubin (2011) argue that the benefits of free services outweigh the costs of identity theft and other risks to which users might be exposed. In their view there is no economic argument based on data for increased privacy. The cost of internet fraud is estimated at $28bn per annum in the US and Europe, but the benefit of free services based on provision of personal data is estimated at $100bn, a substantial economic surplus.

This approach reduces privacy and access concerns to a simple economic argument and requires a lot of assumptions to be made on the value of economic benefit or loss associated with use of personal data. It does not take into account the potential loss of income associated with personal ownership of data. Tene and Polonetsky (2012) focus on data that has been aggregated and anonymised. They argue that the cost of protecting personal data to prevent re-identification would mean that "*many beneficial uses of data would be severely curtailed*". They suggest further research for one strand of this investigation: "*We call for the development of a model where the benefits of data for businesses and researchers are balanced against individual privacy rights*".

DATA PROTECTION

Data protection is a complex concept whose meaning and scope continue to evolve. There are differences in scope between the concept as applied in UK's Data Protection Act 1998 and the EU Data Protection Directive (95/46/EC), which are discussed in Chapter 7. The concept continues to evolve with the General Data Protection Regulation 2012. Data protection is essentially a legislative response to the need for privacy **and** the need to protect individuals against abuse of personal data. Abuse or misuse of personal data does not necessarily mean that privacy is breached. For instance, if personal data gathered for one purpose is used for another, there may be no breach of privacy, but it could still adversely affect the individual concerned. An example would be where one government agency gathers data in order to assess an individual's right to benefits, and another government agency used that data to prosecute the individual. Automated processing of data and large-scale collection of personal data increase the risk of misuse, because of potentially widening access to that data and re-use of that data for purposes beyond those for which it was collected.

Another concern is the introduction of technologies such as flash cookies which regenerate http cookies if they are deleted by the user. This effectively takes away users' control over whether to be tracked or not (Mendel et al. 2012; Sipior et al. 2011). Consent is most commonly manifest as an opt-out, or less commonly, an opt-in feature for tracking and use of personal data with features such as the Network Advertising Initiative's opt-out cookie (Wills & Zeljkovic 2011). If a service changes the defaults, this can cause problems. For instance, when Facebook changed its beacon feature from opt-in to opt-out there was a major outcry and user rebellion which forced Facebook to restore the original opt-in condition (Brunger 2010). The advertising industry's own guidelines recommend that users should give explicit consent before personal data is passed on to third parties (Gray & Mills-Wade 2011).

Ownership and control of personal data is also one of the underpinning principles of the Privacy by Design (PbD) approach (Information Commissioner's Office 2008).

The concept of control or self-determination of personal information has been incorporated into the data protection legislation (Wacks 2010, p.111). The idea of consent has been further developed in the General Data Protection Regulation 2012 to 'informed consent'. This means that an individual has to consent to the uses to which his or her personal data might be put. This particularly applies where there is a change of purpose.

DATA HANDLING BY SOCIAL NETWORK SERVICES

WHO ARE THE PLAYERS?

In order to understand how personal data is used in the context of SNSs, it is necessary to identify the players or agents involved in gathering, distributing and processing that data. Figure 5 shows how personal data and advertising data flows between the different agents.

Users and their contacts (other users) are grouped together as the advertisers may not necessarily distinguish between them. Users provide personal data to their SNS provider via an ISP (Internet Service Provider). The ISP is included because as an agent it may be subject to regulation or to legal action by other agents. The SNS provider may make personal data available to associates and affiliates or to advertisers, who may be affiliated organisations or third parties. The privacy policies of many SNS providers (Chapter 8) refer specifically to sharing personal data with third parties (usually anonymised) or else with affiliates. Previous studies have shown that affiliates can number in the hundreds or even thousands, depending on what definition of affiliate is used. An investigation of the top 50 internet services showed that some providers were part of groups with up to 2,300 subsidiaries (Gomez et al. 2009).

Personal data is also relayed to other users as an activity log ('X has just updated their profile', or 'X has just made friends with Y'), either directly or via groups that they have in common. This may be seen as less of a problem because by becoming a friend or connection with someone there is an implied expectation that personal details will be shared with them and among the groups they belong to.

It becomes more relevant when the SNS provider shares personal data with third parties or with associates and affiliates (in some cases contractors are claimed as affiliates). This may be anonymised data – many SNS privacy policies make reference to this, or it may be identifiable personal data. For this reason they have been grouped together. The

advertisers then push tailored advertisements to targeted users. In doing so they may use tracking technologies to monitor internet behaviour and to build up profiles of individual users. This can be used with a registration system or login to a service provided by the advertising company to create identifiable (i.e. not anonymised) personal data.

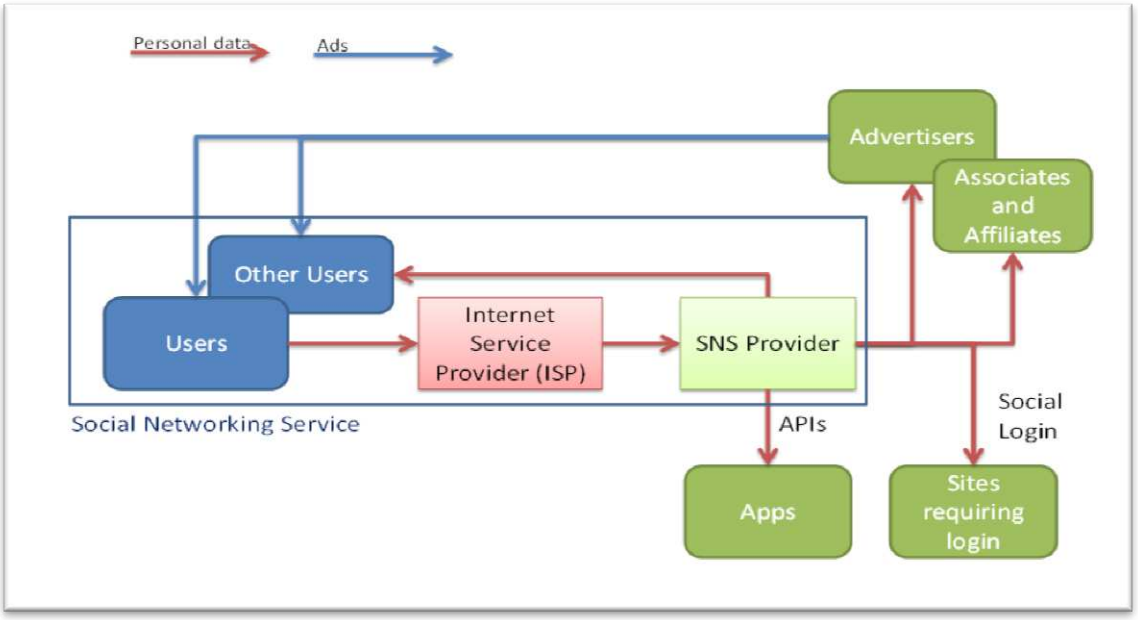

FIGURE 5 - RELATIONSHIP BETWEEN DIFFERENT AGENTS IN AN SNS

INFORMATION MADE AVAILABLE TO THIRD PARTIES

The concept of third parties appears in the literature and in commentaries on the use of personal data in SNSs (Wills & Zeljkovic 2011). This is seen as one of the major risks associated with personal data and is reviewed further in the chapters on risk (Chapter 5) and on privacy policies (Chapter 8). For the purposes of this study, the EASA definition is used. It says that a third party is an organisation that engages in online behavioural advertising on websites that it does not own or operate (Gray & Mills-Wade 2011). This definition is also used in the CAP Code (Committee of Advertising Practice 2014), representing the regulators.

Third parties mostly access personal data by means of tracking technologies such as cookies or web beacons (Pierson & Heyman 2011). Smit, Van Noort and Voorveld (2014) define cookies as *"small text files that are put on users' devices, such as notebooks or smart phones, to facilitate the functionality of a website (first-party, session or functional cookies) or to collect profile information for targeted advertising (third-party or tracking cookies)."* Gomez, Pinnick and Soltani (2009) have already pointed out the contradiction between a commitment to not share personal data with third parties whilst still allowing the use of web beacons. Some

services such as Google have identifiers that are used on mobile devices instead of cookies to track users (CNIL 2012). Other techniques are scraping personal data from discussion forums and device fingerprinting (Christiansen 2011).

Even where there is a commitment not to share personally identifiable data with third parties, it is possible for third party cookies to masquerade as first-party cookies (Krishnamurthy & Wills 2009). This means that personal data is still available to third parties, despite assurances of SNSs.

Many of the SNSs provide varying amounts of metadata about personal profiles and data from those profiles via APIs provided by the SNSs. An apps (applications) developer can create APIs to request standard sets of metadata from SNSs during social login (where an individual uses their social profile username and password to validate their identity and log into a third party website) and can request additional information. For example, JanRain, one of the major commercial providers of identity services offers social login to more than 30 social networks. Facebook provides the following basic personal information to their clients (the third-party website or Apps owner) via JanRain's social login service:

- Address
- Birthday
- Email
- Profile Photo
- Verified Email
- Display Name
- Gender
- Homepage
- Identifier
- Name
- Preferred Username
- UTC Offset

Additional personal information can also be accessed during social login, including personal data about contacts (such as Facebook Friends).

CONCLUSION

Privacy is not co-terminous with data protection. Misuse of personal data (which data protection legislation is meant to prevent) is distinct from, although related to, privacy. Nevertheless it is necessary to consider what is meant by privacy. The literature review covered the origin of some definitions of privacy. For instance, Westin (2003) identifies "*four psychological conditions or states of individual privacy—solitude, intimacy, anonymity, and reserve*". However later commentators have made the distinction between information privacy and other types of privacy such as that related to personal space (Smith et al. 2011).

Is privacy necessarily a good thing? The starting point for many studies is that personal privacy has to be preserved and that any failures in the regulation of access to personal data is necessarily a bad thing. However some commentators caution against blanket application of privacy laws to anonymised (and consolidated) information because "*many beneficial uses of data would be severely curtailed*" (Tene & Polonetsky 2012). The benefits of exchange of personal data, include: improved services to users; cheaper delivery; and better security and protection.

Some research suggests that despite stated concerns about privacy, people often reveal considerable personal data when they start to interact with an online service. This is known as 'the privacy paradox' (Sobel 2007; Bonneau & Preibusch 2009).

Although the concepts of privacy and data protection are closely allied, they are not the same. Privacy is a right identified in Article 8 of the European Convention on Human Rights. However a more sophisticated treatment of privacy is needed beyond Warren and Brandeis's (1890) *"right to be let alone"*. The definitions of privacy covered in this chapter this showed that there was not a single consistent view of privacy. Information privacy has gained ground as a concept and has been found to be a useful focus for attention when considering SNSs.

Privacy is bound up with data protection, but it is important to make the distinction between these two concepts. Defining privacy and the limits to this concept is a starting point. For instance, distinctions can be made between privacy and: non-intrusion, seclusion, secrecy, and autonomy (Tavani 2000). Other commentators have made a distinction between 'personal' and 'private' which have distinct meanings in the context of data protection (McCullagh 2009). Data protection includes commitments to controlling access to personal data and therefore some aspects of information privacy. It also encompasses issues about use of personal data as well as information quality (data accuracy) and appropriateness (relevance to the stated

purpose). Data protection is also one mode of regulation whereas information privacy is an outcome or an objective of this type of regulation.

Krishnamurthy and Wills (2009) discuss personally identifiable information at length and suggest that much of the data gathered from SNSs by advertising aggregators can in fact be 'reverse engineered' to identify and associate specific individuals with this type of data. This throws into question the assertion of some of the SNS providers that this type of information is not 'personal information' and therefore not subject to the same controls and protections as other categories of personal information.

For the purposes of this study the definition of personal data incorporates behavioural data and aggregated personal data, especially that which is gathered by tracking technologies. Despite attempts by SNS providers to make the case for excluding these types of data from data protection considerations, there is a strong case for this to be regulated along with other types of personal data, because ultimately it is possible to identify individuals.

CHAPTER 5 –RISK

INTRODUCTION

The content of this chapter has been published as a peer-reviewed paper (Haynes & Robinson 2015).

BACKGROUND AND CONTEXT

Users of SNSs make personal information available to social network providers in exchange for 'free at the point of use' services. This personal information is voluntarily provided by users, and is usually covered in the Terms and Conditions of Service or is gathered by service providers who track online behaviour using agents such as 'cookies'. Making personal data available to a wide audience exposes users to risk. Although there have been attempts to enumerate some of these risks, which are described below, there has not been a comprehensive review of the risks or any attempt to develop a model of user risk in the context of SNSs. There is a tension about the relative importance of individual and social factors in the study of information behaviour (Bawden & Robinson 2013). This is apparent in the individual response to social media and the way in which different interest groups regulate access to personal data.

An Oxis survey suggested that contrary to popular perceptions, users are becoming more aware of privacy as a concern on the Internet, especially when it comes to using social media (Dutton & Blank 2013). A comprehensive review of Facebook research in the social sciences recognised the need for researchers to analyse the risks associated with Facebook use (Wilson et al. 2012, p.216):

> *By better understanding the threats to privacy, researchers and developers can construct countermeasures to mitigate the risks, and users can take informed steps towards protection their personal information*

This chapter sets out to identify the risks to individual SNS users and to develop a model of risk that can be applied more widely to internet use and social media as they continue to evolve. The following research questions are explored in this chapter:

- What are the risks to individuals that are associated with personal data on SNSs?
- Is there an existing typology of individual risk that adequately covers SNSs?
- Can a model of risks to users be used to differentiate between possible regulatory responses?

Regulation is one area where an up-to-date and relevant model of risk could contribute to improved protection of users. Risk-based regulation has emerged as a dominant approach in Europe and the UK in the last few decades. Baldwin, Cave and Lodge (2012, p.83) suggest that *"Regulation can be seen as being inherently about the control of risks…"*. This is a view supported by Hutter (2006, p.205): *"...regulation has come to be defined as controlling and also as a way of managing risks*".

METHODOLOGY

In order to address these questions, this research was based on a systematic review of the literature, and a survey of information professionals in the UK. Modelling techniques were used to develop a concept of risk that is relevant to internet use and, more specifically, to SNSs. The literature review identified general risk typologies which were analysed in terms of: their applicability to SNSs; their focus on risk to individuals; and their ability to distinguish between types of risk to individuals.

A survey of library and information service (LIS) professionals in 2014 provided insight into the perceived importance of different risk categories (Appendix C). This sector was chosen because it is a well-developed professional group representing users (many LIS staff act as intermediaries), and who are information literate and are therefore likely to be exposed to a wide range of online scenarios. It is also a cohesive group with a track record of active use of social media (Cooke & Hall 2013). The survey was directed at UK users of SNSs using a filter question at the start of the survey to exclude non-UK users. This was cross-checked against the location of the IP Address of the device accessing the survey and logged by SurveyGizmo. The survey objective was to identify the range of risks to which users are exposed and to gain some insight into the perceptions of risk and priorities for managing risk. The survey was based on purposive sampling directed at LIS professionals in the UK, using a variety of forums (listed in Appendix C) to generate a snowball effect (David & Sutton 2011, p.232). Participants were encouraged to publicise the survey through their own professional and personal networks.

A model of risks was developed from an analysis of the consolidated lists of risks identified in the survey and the literature. A typology was developed which formed the basis of a model of personal risk in SNSs. The event and consequence of each risk was analysed to identify the relationship between the risks and to develop a definitive set of outcomes which might have the potential as a tool to evaluate different regulatory approaches.

PRIVACY AND RISK

Information privacy is an important aspect of any discussion about personal data on SNSs. The volume of personal data available on SNSs puts it firmly in the category of 'big data'. It has been suggested that when dealing with big data *"the change of scale leads to a change of state"* and that *"this transformation not only makes protecting privacy much harder, but also presents an entirely new menace: penalties based on propensities"* (Mayer-Schönberger & Cukier 2013, p.151). For instance, where security agencies try to prevent terrorist acts by pre-empting them, individuals are targeted and may be arrested or have their movements restricted without being convicted of any crime. Another problem is 'fetishizing'. This is a common fallacy identified elsewhere (Hansson 2004), where because the picture provided by big data is so compelling, it becomes the over-riding factor in making a decision or judgement.

A Unesco report identified a range of privacy issues associated with the Internet. While these are not expressed as risks they could lead to users being exposed to risks. The issues identified are:

- User identification – unique identifiers, cookies and other forms of user identification
- Adware, spyware and malware conduct covert data logging and surveillance
- Deep packet inspection (DPI)
- Pervasive geo-location technology: an emerging threat to Internet privacy
- Data processing and facial recognition
- Internet surveillance technology

(Mendel et al. 2012, pp.39–49)

Anderson (2013) talks about the difficulty of applying technical 'quick fixes' to complex social systems. This can lead to mismatches between users' expectations and the behaviour of SNSs. He identifies a number of scenarios to illustrate this:

- Attacker re-posted private entries which included sensitive information in a more public forum
- Permissive default privacy settings
- Changes to privacy settings by SNS provider without consent of users. This means that formerly private friends lists are exposed to public view
- Apps developers harvesting personal data to third-party advertisers and data aggregators (in breach of terms of reference)

- Cautious users unwilling to expose themselves to risk and thus being severely limited in what they can do

He goes on to point out that the big differences in power between service providers and users, effectively mean that users have little choice or control over their own data once they sign up to SNSs.

Nissenbaum (2010) identifies three types of privacy issue in social media:

1. Individuals post information about themselves, which later gets them into trouble, with an employer, for instance
2. Posting information about other people, often without their explicit permission can cause problems. Even where there are remedies, such as removing tags from photos, the photos may still remain on the system
3. Harvesting and use of personal data on social networks by advertisers

DEFINING RISK

Risk is an elusive concept based on the notion of uncertainty sometimes expressed in terms of the probability of an adverse event occurring. Commonly-used definitions of risk as *"a situation involving exposure to danger"* or *"the possibility that something unpleasant or unwelcome will happen"* are not very specific and need to be pinned down (Pearsall & Hanks 1999, p.1602). The international standard on risk management, ISO 31000:2009 (British Standards Institution 2010), starts with an even more general definition *"effect of uncertainty on objectives"* and goes on to say that *"An effect is a deviation from the expected – positive and/or negative"*. The Standard does eventually provide a more specific definition: *"Risk is often expressed in terms of a combination of the consequences of an event (including changes in circumstances) and the associated likelihood of occurrence"*. However risk is more widely understood to be an event with a negative outcome, in other words, a threat: "*Risk refers to uncertainty about and severity of the events and consequences (or outcomes) of an activity with respect to something that humans value*" (Aven & Renn 2009, p.2). From the regulatory sphere a working definition is: *"…risk is usually defined as the probability of a particular event (or hazard) occurring and the consequent severity of the impact of that event"* (Baldwin et al. 2012, p.82).

For the purposes of this research risk is defined as an uncertain event which has an adverse impact on an activity or outcome. Applied here, risk is an event of unknown probability of occurrence involving personal data on an SNS that has a negative impact on that person. For

instance, an individual's data might be copied for the purposes of fraud, resulting in that individual suffering financial loss.

TYPOLOGIES OF RISK IN THE LITERATURE

A GENERAL TYPOLOGY OF RISK

Some early commentators have attempted to identify risks associated with the use of SNSs (Rosenblum 2007). However going back to more general approaches to risk identification provides a wider picture. There can be a distinction between physical and social risks which can be integrated (Macgill & Siu 2005, pp.1108–1110). Tulloch (2006, pp.132–133) adopts a social approach to risk:

> *Thus, it seems clear that current research is positively engaged with the construction of self-identities in conditions of risk that these frequently take account of the reflexive concern for dialogic negotiation within and between everyday 'lay voices' and professionals, and that by and large this work* [...] *embeds 'wider social understanding' analysis in quite traditional understandings of the 'otherness' of age, gender, sexual preference, class, and (dis)ability*.

Swedlow and associates' (2009, p.237) research into risk and regulation is based on the *"construction of a universe of nearly 3,000 risks* […] *over a thirty-five year period"*. This provides a comprehensive view of the types of risk that exist generally and is used as a starting point for identifying and categorising the risks faced by SNS users. Some of these risks would arise directly from misuse of data; others are related to the data held about individual history, behaviour and preferences. The following categories from this 'universe' of risks might be applicable to social media and specifically to SNSs:

> **Crime and violence** – There have been a number of court cases where revealing personal data of individuals on social media has exposed them to threats of violence or to harassment (Agate & Ledward 2013)
>
> **Recreation** – A great deal of use of social networks is for recreation rather than professional purposes and it could be argued that the other risks associated with social media fall into this category
>
> **War, security and terrorism** – With the WikiLeaks revelations starting in 2010 and the NSA scandal in 2013 the press has paid particular attention to the security aspects of

personal information (Leigh & Harding 2011; BBC News 2012; Greenwald 2013). The risks to users are two-fold. The first is that identifying information on social networks may be used to victimise or persecute an individual by a state or terrorist organisation. The second is that an individual's identity may be stolen for use by terrorists or by state security agencies and this could expose them to harm

Political, social and financial – Political, social and financial harm can arise from identity theft. For example, if sufficient biometric data is available on a users' profile it may be possible to set up a false identity to gain access to credit or to purchase products with no intention of paying. The individual whose identity has been stolen may be pursued for payment and may even be liable for debts and costs incurred through the fraud

Social risks may include ostracism because of private information being made available inadvertently to a wider audience than intended. For example expression of views that are not compatible with a community's mores (whether it be a religious group, a political party or an ethnically-based group) may lead to some kind of sanction or even expulsion from that group

Human disease / health – Mental health falls under this category. Cases where vulnerable young people have been driven to suicide because of harassment and bullying are an extreme example of this (Wakefield 2014). Less extreme, but nonetheless distressing, may be social isolation and associated depression. Even an affront to an individual's self-esteem and confidence is a potential threat to mental well-being

Occupational – Some employers admit that they search the social media profiles of potential employees and take the results into account in their recruitment decisions (Rosenblum 2007, p.46). It is also an issue for employees who use social media in their private lives to express their views. If an employer deems this to be detrimental to their business or incompatible with their views, it could result in disciplinary action or even dismissal

Consumer products – Consumer products are associated with advertising and this is one of the major areas of concern of many users (Rosenblum 2007, pp.46–47). Behavioural advertising depends on tracking online browsing behaviour and sites visited in order to deduce the interests of the user and target them with advertising

for products that they are likely to be interested in. The impact on users could be described in terms of nuisance caused or possible social isolation

Related risks – A number of the general risks identified are not core to SNS use but may be associated with it in some way. For example, the following would also affect the political, social and financial risks faced by individual users:

- Alcohol, tobacco, and other drugs
- Medication and medical treatment
- Toxic substances
- Human disease / health

In all of these cases the risk is associated with information about these activities being available on personal profiles via social networks. So, for instance, an indication of previous problems with drug abuse may prejudice employment prospects, and health problems revealed online may affect insurance premiums.

OTHER RISK TYPOLOGIES

Other researchers looking at the Internet have provided more relevant categories of risks that might be associated with use of social media (McDonald 2013; Farr 2013; Solovic 2013; Mann 2009). These can be broken down into risk events and associated consequences. Table 3 shows these risks grouped into nine main headings.

TABLE 3 - PERSONAL RISKS ASSOCIATED WITH SNSS

Risk title	Description
EXTERNAL THREATS	
Identity theft	Includes tax-related identity theft. This risk may lead to other consequences such as wrongful arrest or financial loss
Phishing	Fraudulent link or site entices personal information from the user
Malware link	Link to malware (may be embedded in a direct message or attachment) which may result in external monitoring of passwords, or disruption to computer operations
Hijacking of profile	Hijacking of personal site, profile or page could cause embarrassment or inconvenience. Could be a form of bullying as well.
TARGETTING BY OFFICIAL BODIES	
Loss of liberty	Arrest and prosecution for a crime that the user did not commit (identity theft)
Prosecution and recrimination	Prosecution or recrimination for posting offensive comments on social media. Offender's personal data becomes available to the authorities

Risk title	Description
PHYSICAL HARM	
Kidnapping and extortion	Personal information revealing whereabouts, regular travel routes, or activities that leave users open to extortion
Domestic violence	Abusive individuals pursuing former partners
STALKING, HARASSMENT AND CYBERBULLYING	
Cyber-bullying and trolling	Offensive comments made by colleagues – cyber-bullying and victimisation, ostracism, denigration, flaming, trolling
Inappropriate comments by colleagues	Sexual harassment, sexual solicitation
Harassment	Unwanted attention from other users, cyber-stalking, offensive comments, hate campaigns, silent calls, threats from another user
TARGETTING BY CRIMINALS	
Picture of home and possessions shared	Making the user a target for burglars
Home address published	Making the user a target for home invasion
Financial loss	Liability for bills incurred by fraudster (identity theft)
Scams	Often a form of phishing, where the user is required to provide additional personal information (such as bank account details) or where the user is encouraged to send money to the fraudster. This category includes the following scams: dating, work at home, investment, utility, money transfer, weight loss, fake cheques, mystery shopper, debt relief, pay-in-advance credit, lotteries and sweepstakes, miracle cures, imposter, penny auctions, technical support
DISCRIMINATION	
Sharing genetic information	Denial of health or life insurance, Discrimination during recruitment
Loss of opportunities	Refusal of a job or a place at university because of material on a personal profile page
Loss of financial facilities	Refusal of credit or benefits because of information revealed on personal profile. Bad credit rating
WORK RELATED RISKS	
Contravening company policy	Leading to disciplinary action or dismissal
PSYCHOLOGICAL HARM	
Financial records shared	Causing embarrassment with work colleagues and friends
Sharing of genetic information	Invasion of privacy of blood relatives
Release of account details to relatives or executors	Loss of dignity in death. Distress caused to relatives when details not revealed
Loss of privacy	Disclosure of private information
High school pictures shared	Causing embarrassment, doxing, outing
ADVERTISING	
Persistent advertising	Continual, persistent advertising causing nuisance

Risk title	Description
Spam	Unwanted marketing, junk mail, sales calls, text messages, invitations to connect that contain spam pointed on someone's network update, discussion group spam

RISKS IDENTIFIED IN EUROPEAN UNION LEGISLATION

The European Economic and Social Committee (2010) issued an opinion on social networking services, which particularly highlights the risks to children and "*those with poor digital literacy*". It identified the concerns about "*the risks of the illegal and abusive use of SNS, which rides roughshod over a number of basic human rights*." It identified threats to individuals and more generic risks that happen to users of SNSs. Risks that might be relevant in the workplace include:

Cyber-bullying

Privacy breaches

Reputational damage

Assault on personal dignity

RISKS ASSOCIATED WITH GEO-LOCATION DATA

Geo-location data is an increasingly important part of the delivery of SNSs. By allowing their location to be uploaded by mobile service providers and applications providers, users benefit from enhanced services such as location of nearby restaurants, identification of friends in the vicinity and local maps. However there are also concerns about the risks that users are exposed to when their location data is available. This is a problem that the European Commission is well aware of (Article 29 Data Protection WP 2013).

A number of mechanisms by which geo-location data is gathered or can be reconstructed have been identified. These raise some concerns about the resulting loss of privacy (Andrienko & Andrienko 2012). Andrienko and colleagues (2013) go on to enumerate the ways in which geo-location data is gathered:

- Whenever a mobile device is in use it sends a signal to the service provider. However the provider can send a silent text message to force active communication without alerting the user
- Call data records are another source of geo-location data, which came to prominence in the NSA revelations in 2013 and these can give time-based data on movements (Greenwald 2013)
- Signal strength data can be used to triangulate the position of a mobile device

- Users often consent (not always in an informed way) to their location being identified by apps providers or the mobile service provider for enhanced services. This data might be associated with the user ID which has obvious privacy implications
- Anonymous location data seems to provide better protection, although the authors show how identity and even time-based movement data can be reconstructed
- Some non-location data such as accelerometer data, which is freely available from some devices, can be used to deduce the location with a reasonable degree of accuracy

The description of these mechanisms helps to highlight how easy it is for geo-location data to be gathered without the knowledge or understanding of the user, and how this information is available to service providers, mobile operators and apps providers.

Risks Identified in the Survey

The survey of UK-based LIS professionals (Appendix C) ranked risks to provide an indication of priorities. The score is a weighted calculation. In Table 4 the item with the highest score is ranked first. In each case the score is the sum of all weighted rank counts:

Table 4 - Ranking of risks by LIS professionals

Item	Score	Overall Rank
Identity theft	1934	1
Strangers able to see sensitive personal details	1841	2
Targeting by advertisers	1575	3
Victim of fraud	1531	4
Discrimination by employer or potential employer	1443	5
Targeting by criminals (e.g. so that they can burgle your home while you are away)	1411	6
Friends, family or colleagues able to see sensitive personal details	1297	7
Cyber-bullying or harassment (including stalking)	1288	8
Targeting by official bodies or security agencies	980	9
Extortion or blackmail	628	10
Prosecution by authorities because of crime allegations	590	11
Physical violence or kidnapping	451	12

'Identity theft' and 'Strangers being able to see sensitive personal details' both had high scores in the ranking. Identity theft can itself expose users to other risks such as fraud (ranked 4) and one of the consequences can be financial loss. For instance, if a user's identity is used to apply for a loan or credit facilities, the victim may be left with the liability to pay back the loan.

'Strangers being able to see sensitive personal details' ranked much more highly than 'Friends, family and colleagues being able to see sensitive details'. There is a dual risk of strangers seeing personal details – firstly as a means to commit fraud, and secondly because it exposes users to discrimination by potential or actual employers, for instance. Additional comments from users were concerns about reputational damage and loss of face. Personal information may be exposed by the actions of others, such as when friends mention an individual or tag photographs or other entries with their names (Thomas et al. 2010).

Some of the risks may have consequences that are more to do with social awkwardness or annoyance rather than loss of money or physical threat. For instance, targeting by advertisers may be irritating rather than life-threatening. Potentially there is the loss of face if another person makes assumptions about an individual on the basis of advertising that appears on a screen. There is also the inconvenience of screen clutter and slowing down of browsers if there are a lot of graphics or moving images to download.

A Consolidated Model of Risk

Developing a Typology of Risk

Consolidation of these risk categories yields a typology of risk related to use of SNSs. However not all these risks are related to access to personal data, but relate to intellectual property, security and organisational issues. Three approaches to devising a typology of risk for this domain were considered. Risks can be categorised by:

Risk event

Stakeholder affected

Consequence

Risk Event

A risk consists of an event, for which there is a degree of uncertainty about whether it will occur AND the consequence or outcome should it occur. The first part of this definition is the 'risk event'. Risks associated with the use of SNSs can be categorized according to a general set of risks such as those identified by researchers at Duke University and Northern Illinois

University (Swedlow et al. 2009). These are based on risk events or threats. This categorisation does not take into account severity, or impact, or which stakeholders are affected.

Some threats or risks could fall under more than one heading. For instance, identity theft could be under 'Crime and Violence', if it leads to fraud and eventual financial loss to the individual whose data was 'stolen'. It could also be under 'War Security and Terrorism', where identity theft (the same event) results in a different outcome – a terrorist using an alias to escape detection, for instance. It could be argued that this might expose an individual to even greater harm such as the loss of liberty or even loss of life.

STAKEHOLDER AFFECTED

Risks can be analysed in terms of the stakeholders. In a pilot investigation (Appendix B) prior to the survey the SNS stakeholders were identified as: users, service providers, advertisers, employers, and government. However because this study considers the risks associated with allowing access to personal data on SNSs, it is not surprising that the majority of risks will primarily affect users. Indeed a preliminary analysis of the risks identified to date (Table 3) bears this out. Apart from work-related risks which primarily affect employers, the remaining risks all have some direct impact on users.

Although the main risks are faced by users, release of personal data can have a negative impact on employers by damaging reputations or exposing them to legal action or prosecution. There might be wider risks to government or society if personal data is misappropriated and used for terrorist activities or economic sabotage, for instance. Many of the risks to employers of using SNSs in the workplace are not related to access to personal data. They include issues such as: time wasting, security breaches, copyright, and libel where staff members post inappropriate materials on an SNS site during work hours or on a site with a strong presence by or association with the employer.

The other side of the argument is determining who benefits from access to personal data. Advertisers, and those that pay them or whom they pay, benefit directly from accessing personal data, consolidated or not. Indirectly government benefits because of increased tax revenue from the resulting economic activity. Potentially users also benefit – because of more tailored experience of services and targeted advertising – presumably some value is perceived otherwise no-one would follow the links and there would be no point in advertisers using this as a method of gaining new custom.

CONSEQUENCE

The risks identified when the EU's Data Protection Directive (95/46/EC) was being developed can be divided into two categories: tangible risks; and intangible risks (Lynskey 2012):

Tangible risks

- Discrimination
- Identity theft
- Abuse of power by the state
- Physical harm

Intangible risks

- The chilling effect
- The feeling of helplessness
- The apprehension of future harm

This grouping moves towards the idea of categorising risks by their consequences rather than by the nature of the risk event. This can be further refined by concentrating on consequences to users specifically (see Table 5). This provides a means of quantifying the risks, banding them in risk severity categories, or at least a relative ranking.

TABLE 5 - RISK BY CONSEQUENCE TO USER

Consequence	Risk events or threats that leads to the consequence
Self-harm	Cyber bullying Exposure of sensitive personal data to wider view Inappropriate advertising to susceptible individuals or groups
Loss of self-esteem	Cyber bullying Exposure of sensitive personal data to wider view
Social isolation	Cyber bullying Exposure of sensitive personal data to wider view

Consequence	Risk events or threats that leads to the consequence
Financial loss (e.g. job or insurance costs)	ID theft leading to fraud and financial loss Discrimination in employment or during recruitment because of content of SNS profile (e.g. activities, views or past history – membership of a particular group, or health) Higher insurance premiums because of perception of greater risk based on SNS profile (Health, exposure to hazards, risky behaviour) Use of personal data to target for crime – e.g. burglary during holidays or robbery based on recent purchases Cost of inappropriate purchases made under advertising pressure
Loss of liberty – e.g. injustices because of mistaken identity	ID theft leading to mistaken identification as a terrorist Inappropriate use of personal data by security services to profile and target potential terrorists
Violence against the person	Targeting individuals for stalking Using personal data to get at a target for revenge, robbery, stalking (Rosenblum 2007, p.47)
Nuisance	Appropriation of personal data (aggregated or identifiable) by advertisers

Although this is a useful model, one event could lead to several different consequences. For instance, loss of personal data (an event) could lead to harassment (consequence) or fraud (consequence). One consequence could also have several different causes. For example, financial loss could be as a result of following up inappropriate advertising, or it could be because of identity theft, or because of discrimination by prospective employers who have gained access to personal profiles.

A further complication is that a consequence such as cyber-bullying arising from exposure of sensitive data to an inappropriately wide group, could itself lead to further consequences such as self-harm, loss of self-esteem and social isolation.

From the early days of SNSs researchers have identified different standards of behaviour on the internet as a potential source of risk: "*This artificial sense of the anonymity of Net communications leads people to actually lower their inhibitions, and to feel protected from the consequences of their speech*" (Rosenblum 2007, p.45).

DISCUSSION

A RISK MODEL FOR SNSS

Any categorisation is to some extent arbitrary and so it is necessary to identify what criteria are used to select an appropriate approach. Very few commentators in this area have explicitly selected one or other of the three approaches discussed in this paper – analysis by: risk event; stakeholder; or consequence. For the purposes of this study the key consideration is whether this allows differentiation of risks in terms of possible regulatory responses.

Swedlow and colleagues (2009) analysed by risk event using categories that are too general for this study. The majority of relevant risks that they have identified, fall into a single category – Political, social and financial risks. The categories defined do not deal very well with the consequences of risk events such as: harassment; nuisance; loss of dignity; or invasion of privacy.

The stakeholder approach is used by other researchers focusing on risks specifically associated with SNS use from an employer's perspective (Langheinrich & Karjoth 2010). They go beyond the scope of this study by including risks associated with company information as well as general exposure on social networks. However they identify many relevant risks and this coupled with other analyses that focus on the user perspective, results in a list of risks based on stakeholder groups. This offers a method for investigating the effects of regulation (Ellison & Boyd 2013). The same event (e.g. sharing personal data with advertisers) may have quite different effects on each group. For instance, making personal data available to the partners of an SNS provider may be good for advertisers and some consumers, and bad for other users (especially those not looking to purchase).

There are two main problems with the stakeholder approach. The first is that the majority of risks associated with inappropriate access to personal data will directly affect the user. As this study is concerned with risks to individuals, this is not a good way of distinguishing between risks. The other problem is that the list is long and un-differentiated within these two main categories, with overlap and potential gaps in coverage.

The third approach analyses risk in terms of its consequences and this provides a smaller number of main headings under which risks can be grouped (see Table 5). This approach also allows addition of a stakeholder aspect so that analysis by this criterion is also possible.

The survey brought in wider perspectives on what the risks to individuals were and how those risks interacted. Analysis of the risks identified and the relationships between those risks provides a clear distinction between risk events and their consequences. A map of the relationships between risks categories was developed (Figure 6) from the typology based on consequences of risks events (Table 5). This allows the development of a model of risk relationships. The model emphasises the difficulty of defining limits around the definitions of each risk category, a pre-requisite for measuring or quantifying risk.

The analysis of consequences produces a more complex picture than a simple listing (Table 3) can reveal. One of the challenges of trying to analyse risk is that some consequences may themselves expose individual to new risks and therefore to other types of harm. The figure uses red arrows to point to the risk consequences and labelled black arrows to look at the relationship between underlying risks.

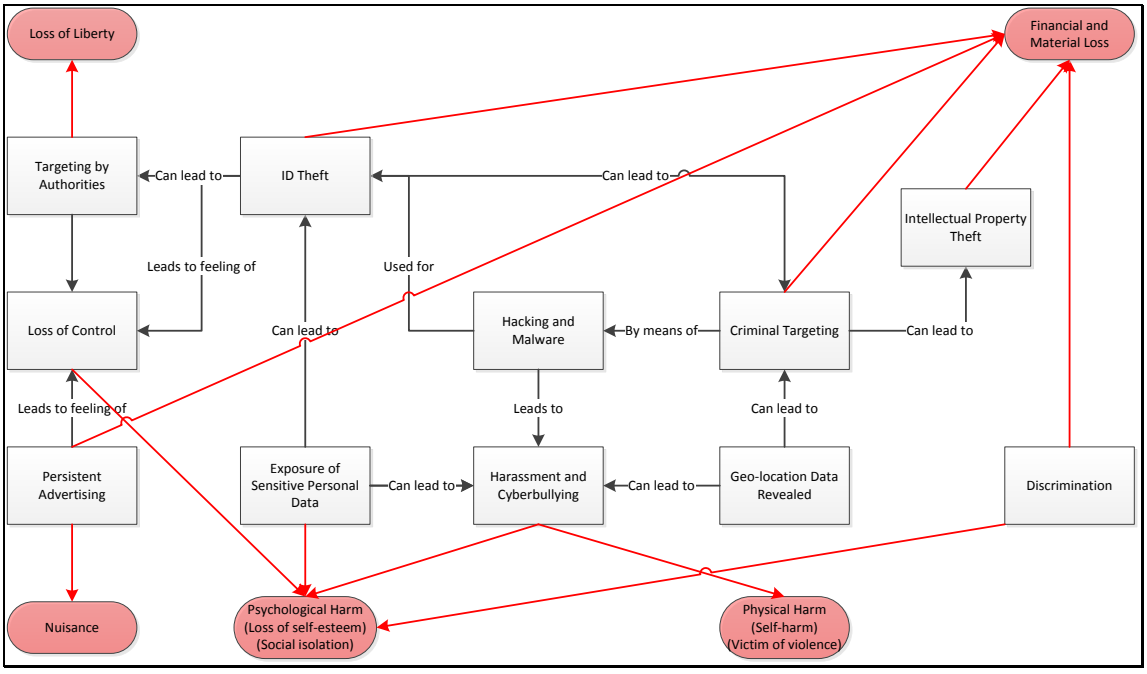

FIGURE 6 - RELATIONSHIPS BETWEEN RISKS AND CONSEQUENCES

This grouping of risks has allowed an inductive derivation of five categories of consequences to users. Within each category, the contributing risks events are described.

Nuisance includes being bombarded with advertisements or users being inconvenienced by having to go through extra steps to preserve their privacy. This could also include intrusion into private lives by strangers, where no other direct harm is felt.

Psychological harm can result from exposure of private information and also from harassment and cyberbullying. This can range from mild social embarrassment when personal information is circulated to those that the data subject would not be comfortable with, through to victimisation and threats. It can also result from a feeling of helplessness engendered by loss of control over who has access to personal data.

Financial and material loss can arise from criminal targeting through or from fraud as a result of ID theft. Active discrimination in the job market – for instance by religion, race, trade union activity or sexuality, all of which may be inadvertently revealed on SNS profiles. Theft of intellectual property via SNSs – especially where users are encouraged to post pictures, videos etc. could result in loss of revenue (Rosenblum 2007, p.46). There have also been cases reported in the press of people inadvertently advertising when they are away, making them targets for burglary or home invasions (Roberts 2010; BBC News 2013a).

Loss of liberty is a dramatic consequence of personal data being made available on SNSs. This could be either as a result of exposure of criminal activity or being mistakenly identified as a criminal or terrorist (Strauß & Nentwich 2013). Boasts about drug-taking on SNSs or postings about location could be used as evidence of criminal activity. Profiling by security services and police are approximate tools that have led to targeting of innocent people with consequent loss of liberty, political persecution and financial loss.

Physical harm can be a consequence of criminal targeting – for instance during a robbery or a kidnapping. Personal data can reveal information about movements, routines and intent and therefore make it easier for criminals to target the individual. There are also concerns about personal information revealing the location of shelters for those escaping domestic abuse.

CONCLUSION

This research has identified risks that individual users of SNSs face as a result of revealing personal data on their profiles or through their online behaviour. Previous attempts to categorise risk have been too general to adequately describe the risk exposure of SNS users. Where there has been a focus on the risks associated with use of the internet or social media, they have tended to focus on a few specific aspects that were topical at the time. A consolidated list of risks reflected the perceptions of risk among a group of library and information professionals surveyed in the UK.

A list of risks does not, however, describe the relationship between different risk categories. This is important because of the strong interdependence between them.

A risk model that more accurately represents the potential threats to users and the consequences can be used as a tool for investigating different modalities of regulation. As much of current regulatory activity is risk-based, this approach could provide a means of evaluating different regulatory approaches. For example, it might be possible to consider whether proposed changes in legislation tend to increase or reduce each of the risk categories in terms of probability of occurrence and severity of impact.

This up-to-date perspective on user risk is of potential utility to policy makers and decision makers. Legislators need a more nuanced tool than currently exists for evaluating proposed new laws or regulations. Service providers can consider the effect of different privacy settings and proposed new services on users, and systems designers have a tool that they can adopt to demonstrate that they are following 'privacy-by-design' principles.

The risk model also provides a conceptual framework for trainers, educators and information intermediaries. These are all roles that are increasingly forming a part of the role of library and information service (LIS) professionals. Their role in modifying user behaviour by example and by user education could have a significant effect in helping users to derive the greatest benefit safely from SNSs and from social media generally.

SECTION III. REGULATION

CHAPTER 6 – A REGULATORY MODEL

INTRODUCTION

This chapter sets out to review Lessig's model of internet regulation and to see how it might apply to regulation of access to personal data on social networking services (SNSs). The shortcomings in the Lessig (2006) model are addressed by developing a modified model of regulation that incorporates self-regulation, one of the main ways of regulating SNSs.

Baldwin, Cave and Lodge's (2012, pp.2–3) definition of regulation is significant in acknowledging that it goes beyond *"control exercised by a public agency over activities that are valued by a community"*. In the last of the three definitions below Baldwin, Cave and Lodge state *"that regulation may be carried out not merely by state institutions but by a host of other bodies, including corporations, self-regulators, professional or trade bodies, and voluntary organizations."* Regulation can be seen:

> *As a specific set of commands* …
>
> *As deliberate state influence* …
>
> *As all forms of social or economic influence* …

This research started by considering the interaction between risk and regulation. In order to do this a conceptual model has been proposed to describe this interaction. It can be argued that regulation has three main purposes:

- To reduce risk
- To minimise market inefficiencies
- To create opportunities for service innovation

This research focuses on the first of these and proposes to use risk assessment as a means of evaluating the relative effectiveness of different modes of regulation. The model focuses specifically on reduction of risk to users and looks at the nature of the risks that users face. This can be characterised by the degree of personalisation of the data and its sensitivity in terms of perceived or actual harm that might arise from misuse.

A number of fundamental issues have been identified when it comes to risk regulation *"ranging from the definition and identification of risk, critical debates about the principles inherent in any regulatory activity, to fundamental questions on the appropriate institutions for risk regulation"* (Baldwin et al. 2012, p.102). Given the complex nature of risk regulation, it is

important that: *"…any attempt at regulating risks should involve the pluralization of analytical perspectives, rather than the reliance on any one analytical device alone"* (Baldwin et al. 2012, p.102).

RISK TO USERS

Other aspects of protection of personal data, apart from privacy, are protection from abuses such as: fraud, bullying, and harassment. Some of these could be seen as aspects of privacy "*the right to be let alone*", or opening people up to risks such as financial loss through fraud (Warren & Brandeis 1890). This study also includes exposure to risk when personal data is shared with advertisers.

Personal data can be described in terms of proximity and sensitivity of data. Different modes of regulation could be evaluated in terms of its effect on each category of personal data.

Chapter 5 suggested that risks associated with personal data can be described in terms of:

Risk event
Stakeholder affected
Consequence

So, for instance, a risk might be the following event: non-attributable personal data is made available to advertisers via a web beacon or browser cookie. The consequence of this is intrusive advertising, but there may also be a problem of loss of dignity (as when colleagues have sight of advertisements for personal products) or breach of confidentiality (e.g. when buying a present for a friend or relative in the same household or workplace, and they see the ad).

Harassment, bullying and the associated loss of dignity is a significant risk faced by users or by the subjects of users' postings on their profiles. In *Teggart v TeleTech UK Ltd* an industrial tribunal found in favour of the respondent defending their dismissal of an employee for gross misconduct after he posted salacious and damaging allegations about a co-worker on his Facebook profile. Interestingly the tribunal also referred to Article 8 of the Human Rights Act 1998:

> *"When the claimant put his comments on his Facebook pages, to which members of the public could have access, he abandoned any right to consider*

his comments as being private and therefore he cannot seek to rely on Article 8 to protect his right to make those comments."

MODEL OF PERSONAL DATA

In the review of the privacy policies of eleven leading SNSs (Chapter 8), the following types of personal data were identified:

TABLE 6 - TYPES OF PERSONAL DATA

Type of personal data	Data elements
Identity	Name [IDs such as passport number, driving licence number, National Insurance number, NHS Number, Pupil Number] Gender Sexuality Race / ethnic origin Place of birth Nationality Age
Location	Address Current location IP Address
Security	Username User login / password
Attitudes and interests	Interests Religion Political affiliation Holidays / places visited
Education and employment	Occupation / employment status Education Criminal history Schooling / education Scholastic achievement grades Employment history
Finance	Banking details Income Home ownership
Health	Health status Medical history
Personal network	Marital status Family status Personal details of associates (colleagues, friends, household members, relatives)
Behaviour	Activities (browsing history - sites visited, groups joined, services / products purchased)
Device information	IP Address Browser type and version Device ID

These categories of information can be grouped in different ways. For instance the Data Protection Act 1998 makes the distinction between personal data and sensitive personal data. This is primarily so that sensitive personal data can have additional protections.

WHAT IS BEING REGULATED?

Regulation can be viewed in terms of who is being regulated. For instance, is it the industry, their agents, or the consumers that are being regulated? The Data Protection Act 1998 focuses on the responsibilities of the data controller who can in some cases be seen as representing the SNS provider (see Chapter 7 for a more detailed discussion of the legislation). Part of the problem arises in the definition of data controller. Self-regulation is focused on the SNS providers and to some extent the ISPs as manifest in their privacy policies and End User Licence Agreements (EULAs).

Code is normally enacted by SNS providers and their agents when setting up and modifying their services. They can build privacy into the design of their systems so that they default to non-disclosure. However this may be seen as being in conflict with their desire to extend their membership and the range of services available to members.

Mode is where users regulate their behaviour either collectively (as in market demand) or individually by the way in which they interact with services and the degree to which they reveal personal data. User education is seen as one way to improve individual security.

It could also be argued that activities are being regulated rather than individuals and organisations. For instance, exchange and use of personal data could be subject to self-regulation (in privacy policies), legislation (as with the Data Protection Act 1998) or by code (as with data encryption to protect against unauthorised access to personal data).

THE NATURE OF REGULATION

The emphasis of privacy regulation in Europe has moved towards a risk-based approach (Chapter 5). However there are countervailing views that must be considered:

> *"Risk management is in many respects contrary to the essence of capitalist accumulation, which intrinsically encompasses risk-taking. Risk management and risk regulation can be seen as ways of diverting attention from system weaknesses and reducing the status of regulatory activities to little more than 'blaming mechanisms'. The over-emphasis on regulation and risk management*

can make regulatory systems unwieldy and ultimately self-defeating." (Beck 1992, p.156)

If the sole consideration of regulation is to enhance and protect capitalist endeavour, then regulation can be seen as counter-productive. However most societies subscribe to other values such as communal or societal good as well as individual human rights.

LESSIG'S MODEL OF INTERNET REGULATION

Lessig (2006, p.123) has identified four modalities of regulation that affect the internet (Figure 7).

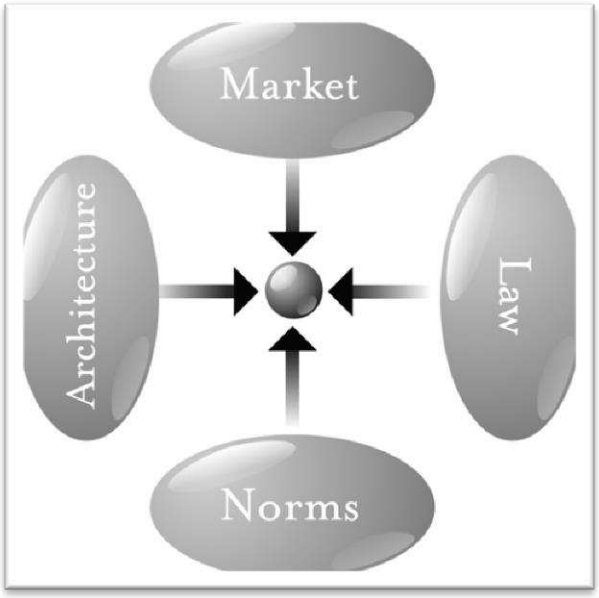

FIGURE 7 - LESSIG'S MODALITIES OF INTERNET REGULATION

Although Lessig's model has become a major reference point of analysis of different regulatory modes, other technology commentators such as Schmidt and Cohen (2013, p.66) have described alternatives based on: corporate, legal, societal and personal responses to privacy and security needs. Each of Lessig's regulatory modes is considered in turn before a new model is proposed.

LAW

A review of the legislation in the UK as it applies to use of personal data on SNSs suggests that the Data Protection 1998 is the main focus for legislative regulation. This in turn is based on the EU's Data Protection Directive (95/46/EC) and associated legislation. It is also dependent on the Human Rights Act 1998, which requires the right to privacy (Figure 8) and the

Communications Act 2003, which regulates some aspects of data communications via the internet.

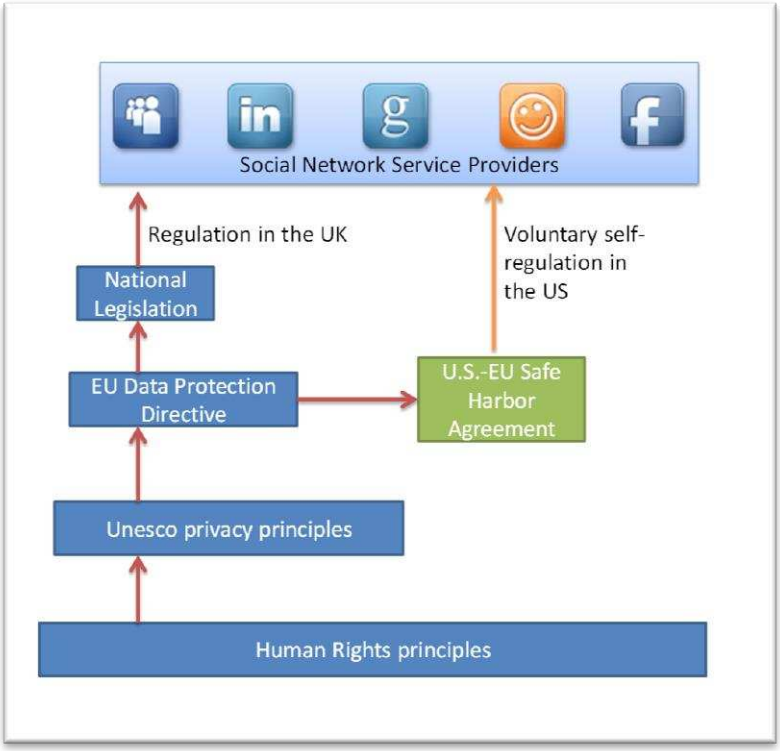

FIGURE 8 - LEGISLATIVE REGULATION IN THE UK

One objection to the application of the Data Protection Act 1998 to SNS providers based in the United States is that they are covered by the U.S.-EU Safe Harbor framework. As far back as 2002 commentators were concerned about the fundamental differences in approach to data protection in Europe and the United States (Muir & Oppenheim 2002). Safe Harbor, as a self-regulatory scheme has been considered weak and with many loopholes that has not been rigorously enforced by the Federal Trade Commission (Connolly 2008; Reay et al. 2009). However a recent study by Solove and Hartzog (2014, p.679) suggest that:

> *Through a gradual process akin to that of common law, the FTC has developed a federal body of privacy law, the closest thing the United States has to omnibus privacy regulation.*

Lessig (2006) does not make the distinction between law and self-regulation. Even when considering legislative frameworks such as the data protection act or advertising and consumer law in the UK, it is difficult to avoid self-regulation as a major component of the regulatory landscape. It also does not take account of the self-imposed commitments that

SNS providers make in their privacy agreements and terms and conditions of service. It could be argued that privacy policies are a type of 'code' in that they reflect the architecture of a system, however they are also an agreement between user and provider, and may be partially covered by contract law.

NORMS

Lessig (2006) describes situations where users contravene accepted behaviour standards and are ostracised online. Users' behaviour and expectations can serve as a powerful regulatory force. This can be applied to SNS and specifically to regulation of access to personal data on SNSs. As part of this research a survey of LIS professionals in the UK suggested that users should have some responsibility for their own online safety (Appendix C). This is a theme that has been picked up by the Information Commissioner's Office. Reports in the press indicated that users' expectations have a powerful effect on SNS providers. Cases such as the Facebook beacon device led to a strong reaction from users that resulted in the withdrawal of the beacon feature because of its intrusive nature (Story & Stone 2007). The feature automatically tracked purchases made online and effectively publicised this information. Boyd (2010) also suggests that norms operate within groups of connected individuals and that certain types of personal information are not revealed beyond the group. Wider social norms about abuse on social media have been covered prominently in the press and this suggests that there are implied standards of acceptable behaviour which when contravened elicit a strong response beyond that of the law (BBC News 2013b).

All this points to the idea that social norms are an important factor in regulating SNSs. The effects of individual behaviour may be governed by collectively-held views of acceptable behaviour, but is only noticed at a market level when large numbers of users respond to breaches of norms.

ARCHITECTURE OR 'CODE'

The way in which systems are designed and the options presented to users are an expression of 'code'. This is about the systems architecture and the way in which it controls access to personal data. At the first level is the amount of personal data gathered by SNSs. This varies significantly – SNSs such as Facebook and LinkedIn offer the opportunity to share very detailed personal information whereas services such as Twitter work on the basis of a minimal profile (up to 160 characters long). The SNSs also vary in the minimum amount of personal data required to register and whether real names or aliases are acceptable. This has led to controversy (described in Chapter 2) where networks such as Google+ have tried to impose the

'real names' requirements on users or where Facebook initially refused to accept preferred names for transvestites or transitioning transsexuals (Boyd 2012; Lee 2014). At the third level is the range of privacy options or settings offered by SNSs. An exploration of privacy policies (Chapter 8) reveals considerable variation in approach. Most operate an opt-out approach, because they want the default to be for greater disclosure – this facilitates connections and provides richer data to sell to advertisers. There is also an argument that it is difficult for users to exercise informed consent in such a complex area and so they will tend to keep the default. However privacy policies may not be an entirely reliable indicator of what options are delivered to users. Firstly, there may be a time-lag between a change to the privacy settings and the published privacy policy. Secondly, the SNS may not keep track of the promises that it makes in its policies and so be out of step with the actuality. This is difficult to verify because these discrepancies may not be immediately apparent.

The idea of privacy by design is intended to encourage developers to consider privacy as a key consideration when designing systems and developing new features. Some regulators have taken this on board and this has become a feature of the proposed European Data Protection Regulation.

Lessig's (2006) model of 'Code' can be extended to include the architecture of the networks and communications services as well as the electronic ecosystem within which they exist. This means that technology-based privacy and identity protection software as well as anti-spyware software are part of regulation by code. As well as the SNSs themselves, there are independent solutions that are available as add-ins to browsers, for instance, to supress cookies and to isolate Trojans and other software designed to capture sensitive data such as passwords and account details. This wider view of code as a regulatory instrument is important when considering the different agents involved.

MARKETS

While there is considerable literature about regulation of the markets, there is little said about the use of markets to regulate an industry. Market effects are recognised, but the emphasis of regulatory bodies is to introduce rules that make the markets operate in a way that the regulators consider desirable.

The growth of internet communities and interactions has allowed the effective development market-driven regulation. This can be seen in social pressure on providers to comply with market expectations. For instance recent changes in the Facebook privacy settings without

full consultation with users led to an outcry and pressure to retract (BBC News 2011). However, this may be because of the implicit threat that legislators may respond with new regulations to address the concerns of their voters. Lessig's (2006) treatment of the market can be seen as a manifestation of user norms. If a sufficient number of users feel strongly about a service and they respond, they operate as a market. It is logical therefore to treat norms and market forces as part of the same regulatory mode which differ primarily in scale.

A Revised Model of Regulation of Personal Data on SNSs

A preliminary survey of users and data protection officers (Appendix B) suggested that there is some scepticism about the effectiveness of the Data Protection Act as a means of regulating access to personal data on social networks. This was borne out by interviews with regulators and industry experts (Appendix D). Although some respondents found it to be effective, many considered that the legislation alone was insufficient. Several respondents saw other modes such as self-regulation, user education, and technology as important elements in the protection of personal data. This can be represented by a model comprising four different modes of regulation:

1. Legislation
2. Self-regulation
3. Code
4. Norms

The validity of this model was tested in consultation with industry experts and through a survey of LIS professionals. It has been developed in the context of SNSs.

Legislation

Legislation corresponds to Lessig's (2006) 'Law'. This includes rules with statutory weight behind them and often involves a national or regional authority which enforces the rules. In the UK the main legislation that regulates access to personal data is the Data Protection Act 1998 and associated statutory instruments. It is based on EU legislation, specifically the Data Protection Directive (95/46/EC), which at the time of writing was due to be superseded by the General Data Protection Regulation 2012. The other significant area of legislation is the Consumer Protection Act 1987 which establishes a self-regulating framework for the advertising industry and affects the digital advertising companies that operate in association with SNSs.

SELF-REGULATION

The Lessig (2006) model does not specifically include self-regulation, and it has been left to others to suggest that self-regulation of the internet is a form of 'norm-based governance' (Cooke 2004, p.36). Self-regulation by an industry can apply if there are suitable sanctions for non-compliance such as expulsion from a group with consequent loss of credibility and market share.

Surrogate regulation, where the responsibility for regulating a professional group or industry is vested in a professional or trade body, can also be effective. Membership of the body becomes a condition of being allowed to trade. This can be seen with the established professions and some sectors in the UK (such as civil engineers, lawyers, doctors and architects). This approach takes the burden of regulation away from the state and the costs of regulation are borne by the regulated individuals or industry.

CODE

Lessig (2006) deals with technical architecture as an instrument of regulation. This is an approach that has been taken up by a number of regulatory authorities, initially in Canada and latterly in the UK and the EU. Code also includes other technology based solutions for managing user identities online or for blocking ads and cookies so that online behaviour is not actively tracked by someone else.

NORMS

This category covers both individual behaviour and attitudes and collective behaviour as seen in markets and covered by the 'markets' category in Lessig (2006). The justification for merging these categories is that regulation by the market is a manifestation of Norms expressed collectively.

CONCLUSION

Figure 9 shows how these different modes of regulation interact with each other and with access to personal data on SNSs.

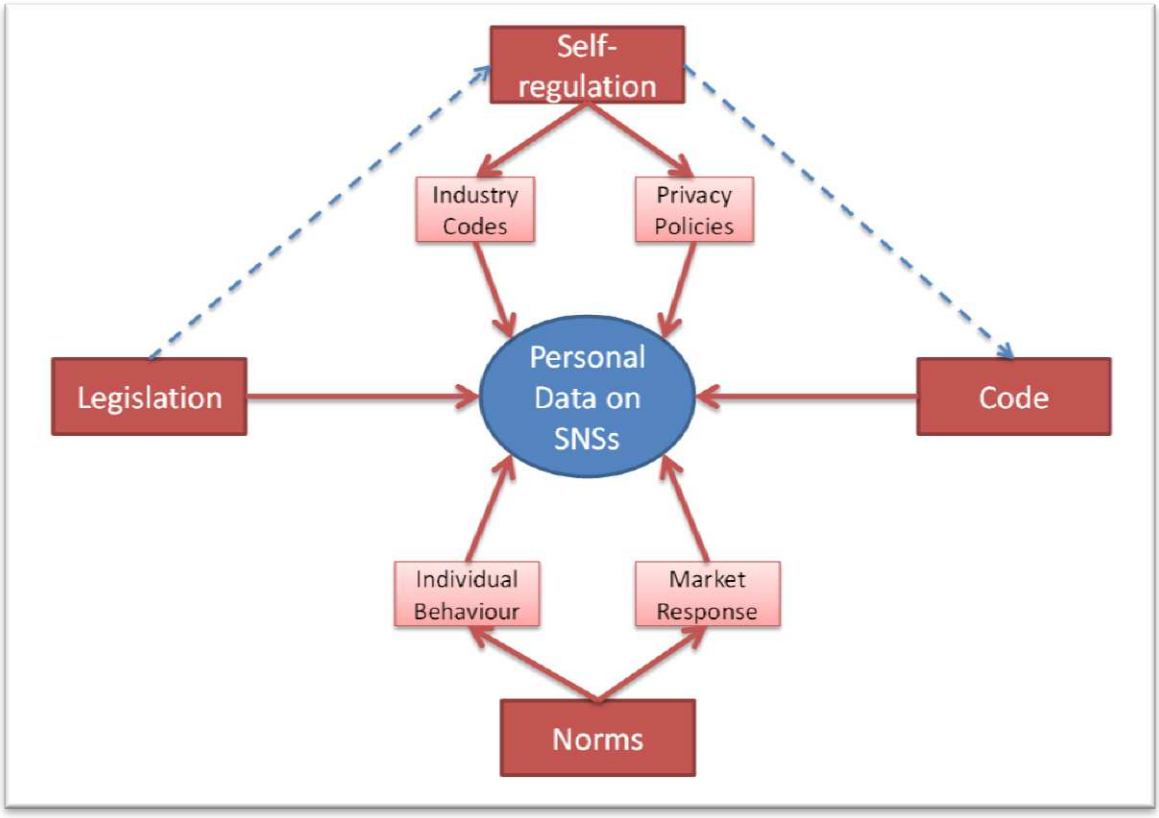

FIGURE 9 - REGULATING ACCESS TO PERSONAL DATA - NEW MODEL

This proposed model of regulation builds on Lessig's (2006) idea that there are four modalities for regulating the Internet. It attempts to cover a major omission in the lack of a category for self-regulation. The model also makes a stronger connection between 'Norms' and collective user behaviour which Lessig treats separately in 'Markets'. This new model also recognises that legislation affects self-regulation which is usually manifest in privacy policies and in industry codes of practice. These codes of practice may themselves by governed by legislation or they may be industry-driven. The way in which enterprises design and deliver SNSs (Code) is itself a form of self-regulation.

CHAPTER 7 – LEGISLATION

INTRODUCTION

Legislation is one of the modes of regulation previously defined (Lessig 2006). In the UK the legislative framework is dominated by the Data Protection Act 1998 (DPA). However other primary legislation and statutory instruments also regulate use of personal data in SNSs. As a member state, the United Kingdom is subject to European Union legislation, which is explored in this chapter. The current legislation is reviewed and this provides a basis for the discussion of the data protection principles. Development of a draft EU Regulation and its implications for future protection of personal data in the UK is also discussed.

THE RULES AND THEIR ORIGINS

EUROPEAN LEGISLATION

UK legislation is governed by the Treaty on European Union 1992 and the EC Accession Treaty 1972. Treaty-based law does not require additional domestic legislation. Voluntary participation by states usually involves the following steps: negotiation, signature, ratification, adhesion, accession. However states can record reservations. Article 21 of the Vienna Convention on the Law of Treaties 1969 allows states to make reservations (i.e. exemptions to parts of a treaty). Other bodies have their own regulations such as European law originating from the European Union. The EU tends not to allow reservations and Article 288 of the Treaty on the Functioning of the European Union 2007 defines what is binding under EU law.

The principle of subsidiarity means that dualist member states such as the UK enact their own legislation within a framework determined by the European Union. European Directives are incorporated into UK national law by means of new legislation, such as the Data Protection Act 1998, or by means of statutory instruments such as the Privacy and Electronic Communications (EC Directive) Regulations issued in 2003, 2004 and 2011 and based on the Directive on Privacy and Electronic Communications (2002/58/EC), also known as the ePrivacy Directive.

European Regulations, as opposed to European Directives, are directly applicable to all member states and are not legislated for nationally. Existing regulations governing competition, discrimination and free movement are examples of this and the General Data Protection Regulation 2012 will fall into this category when it has been finalised.

Laws derived from the European Convention on Human Rights 1950 affect data protection. Unlike the EU, the Council of Europe (a non-EU body) allows reservations (Article 57 of the Convention).

The European Economic and Social Committee (2010) opinion identified risks to users of SNS, referred to in Chapter 6. As well as hazards associated with geo-tagging, and facial recognition technologies, spreading of viruses via SNS was also identified. The Opinion goes on to recommend measures to improve digital literacy. It suggests that SNS providers should self-regulate or participate in co-regulation at Community or national level. This would allow operators to sign up voluntarily to a code of practice that would be monitored and enforced by the regulatory authorities. It also recommends the appointment of:

> *a community-level Ombudsman responsible for all issues relating to the protection of human dignity, privacy and data protection in the electronic communications and audiovisual sectors, with specific responsibility for SNS*.

DATA PROTECTION DIRECTIVE (DPD)

The EU Data Protection Directive (95/46/EC) is binding on the United Kingdom under the provisions of the Treaty on European Union 1992 (also known as the Maastricht Treaty). The DPD is intended to allow "*the free movement of goods, persons, services and capital*" between Member States and "*also that the fundamental rights of individuals should be safeguarded."* The Directive refers in several places to "*the right to privacy*" as one of the rights and freedoms of individuals.

DATA PROTECTION ACT 1998 (DPA)

The DPA sets up the "*provision for the regulation of the processing of information relating to individuals*" and encompasses the eight Data Protection Principles described in the Directive. These in turn are derived from the OECD Guidelines (OECD 1980). These Guidelines arose from a concern about discrepancies between the privacy laws of different OECD countries which could act as a potential trade barrier by inhibiting the exchange of data between countries. The rapid growth of databanks containing personal data was a particular concern.

Section 6 of the DPA defines the role of the Information Commissioner, appointed by Her Majesty and gives him or her specific powers under the Act. Section 6 also makes provision for appointment of members of an Information Tribunal to hear and determine appeals against a notice from the Information Commissioner. Since 2010 the Information Tribunal has become the Information Rights Tribunal in the General Regulatory Chamber of the First Tier Tribunal. This was formed from the incorporation of a number of tribunals into a centralised tribunal system under the Transfer of Tribunal Functions Order (SI 2010/22).

Secondary legislation arising from the DPA includes statutory instruments and case law that sets precedents in the way in which the law is interpreted. In addition the Information Commissioner's Office (ICO) from time to time issues regulations and guidelines which are available on the ICO website.

COMMUNICATIONS ACT 2003

The Office of Communications (Ofcom), set up by the Communications Act 2003, does not have a specific interest in regulating privacy, but does have a role in licensing communications service providers. It also has an interest in attitudes to social media which has been the subject of its own research (Ofcom 2008). Section 127 of the Act prohibits use of public electronic communications networks for sending a message:

> *that is grossly offensive or of an indecent, obscene or menacing character*

or

> *for the purpose of causing annoyance, inconvenience or needless anxiety to another*

This could cover use of SNSs to bully or harass users or for persistent advertising.

REGULATION OF INVESTIGATORY POWERS ACT 2000 AND THE COMPUTER MISUSE ACT 1990

Regulation of Investigatory Powers Act 2000 prohibits unlawful interception of communications including electronic communications:

> *(1)It shall be an offence for a person intentionally and without lawful authority to intercept, at any place in the United Kingdom, any communication in the course of its transmission by means of: (a) a public postal service; or (b) a public telecommunication system.*

This could be applied to personal data on social networks where there is an intent to gain unauthorised access. This is also covered by Section 1 of the Computer Misuse Act 1990:

> *(1)A person is guilty of an offence if— (a) he causes a computer to perform any function with intent to secure access to any program or data held in any computer; (b) the access he intends to secure is unauthorised; and (c) he knows at the time when he causes the computer to perform the function that that is the case.*

Effectively this outlaws hacking into user accounts on social media. However, it could be argued that by signing up to a service and voluntarily putting personal data on an SNS site a user is allowing the SNS provider to pass on personal data to third parties and that this therefore does not constitute computer misuse within the terms of the Act. The applicability of the Computer Misuse Act 1990 depends on 'unauthorised access'. If a data subject gives consent for his or her personal data to be made available via the SNS, it is effectively authorised by them. Many of the concerns that do arise with personal data are about misuse of the data, not misuse of the computer to gain access to the data.

HUMAN RIGHTS ACT 1998 (HRA)

Schedule 1, Article 8 of the Human Rights Act 1998 identifies that:

> *Everyone has the right to respect for his private and family life, his home and his correspondence.*

The UK ratified the Convention for the Protection of Human Rights and Fundamental Freedoms in 1951 and enacted the Human Rights Act in 1998, which came into force in October 2000. It "*gives effect to the rights and freedoms guaranteed under the European Convention on Human Rights*".

A Court (defined in the Act) may make a Declaration of Incompatibility if any provisions of a piece of primary legislation are incompatible with any rights under the Convention. This constitutional role means that any primary legislation (whether it has already been enacted or is proposed in a bill) must observe the right to respect for private and family life. The HRA allows individuals to refer their cases to the European Court of Human Rights once all appeals are exhausted in the UK Court system.

As well as the constitutional role of the HRA, there is the right to privacy which underpins much of the DPA and the presumed rights of UK residents using social media. However the HRA also enshrines the right to freedom of expression. This could apply to online behavioural advertising as well as self-promotion.

SECONDARY LEGISLATION

In the UK statutory instruments are the main source of secondary legislation. These take the form of regulations[2] or statutory instruments. The Privacy and Electronic Communications (EC Directive) Regulations are based on the European ePrivacy Directives.

In addition to the directives the European Data Protection Supervisor (EDPS) issues Notices and Opinions, which provide additional guidance to European Union institutions, national governments of member countries and citizens of Europe.

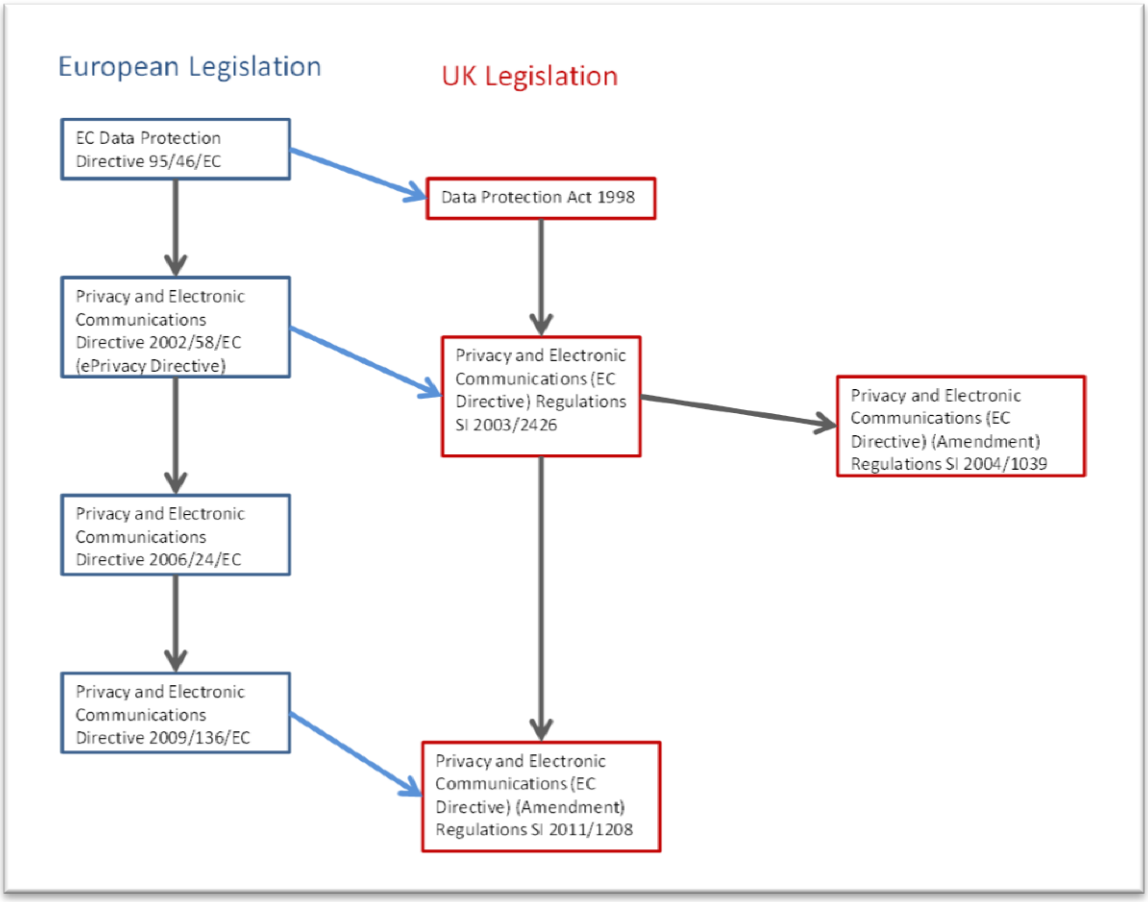

FIGURE 10 - RELATIONSHIP BETWEEN EU AND UK DATA PROTECTION LEGISLATION

Figure 10 illustrates the relationship between European Union and UK legislation. Statutory instruments that have come into effect since the DPA can be applied directly to SNS providers. The Privacy and Electronic Communications (EC Directive) (Amendment) Regulations (SI 2011

[2] UK regulations are not to be confused with European Union Regulations which apply directly to all member states

/1208) gives a new definition of "*personal data breaches*" and places the obligation on data controllers to notify the Information Commissioner of any data breaches. The Statutory Instrument makes provision for penalties for failure to do so and specifies that the Information Commissioner should be notified in the event of a personal data breach. It is an amendment of the Privacy and Electronic Communications (EC Directive) Regulations (SI 2003/2426) and implements Articles 2 and 3 of Privacy and Electronic Communications Directive (2009/136/EC).

The other relevant development for SNS providers is the rationalisation of registration fees for data controllers into two tiers. The Data Protection (Notification and Notification Fees) (Amendment) Regulations (SI 2009/1677) sets the fees payable by data controllers for registration under the DPA at £35 for all charities, companies with a turnover of less than £25.9m and fewer than 250 staff, and for public authorities with fewer than 250 staff. The remainder (so-called "tier 2" data controllers) pay a fee of £500. This might have an effect if any of the SNS providers decide to submit to UK law.

GENERAL DATA PROTECTION REGULATION 2012 (DPR)

The European Parliament (2009) has made a recommendation on strengthening security and fundamental freedoms on the internet. This includes:

- Full and safe access to the internet for all
- Strong commitment to combating cybercrime
- Constant attention to the absolute protection and enhanced promotion of fundamental freedoms on the internet; and
- International undertakings

Article 29 of the Data Protection Directive (95/46/EC) makes provision for the establishment of a Working Party, with representatives of the national data protection authorities of Member States, plus the EU Data Protection Officer and a representative of the European Commission. As well as an annual report it commissions its own research and consultations and advises the Commission on legislative and other measures that can be taken to improve the data protection framework. The proposed DPR will replace the Article 29 Working Party with the European Data Protection Board.

In 2010 the Article 29 Working Party reviewed the framework for data protection in light of changes in technology and emerging practice internationally. The European Commission

(2010) issued a communication that encompasses legislative and non-legislative measures for data protection. It signals specific intended actions by the EC in a number of areas including:

- Notification of breaches
- Data minimisation
- Consent
- Measures for self-regulation and other non-legislative approaches
- Harmonisation of rules and processes

Of particular note are the areas where individual rights could be improved:

> *The Commission will therefore examine ways of:*
>
> *- strengthening the principle of data minimisation;*
>
> *- improving the modalities for the actual exercise of the rights of access, rectification, erasure or blocking of data (e.g., by introducing deadlines for responding to individuals' requests, by allowing the exercise of rights by electronic means or by providing that right of access should be ensured free of charge as a principle);*
>
> *- clarifying the so-called 'right to be forgotten', i.e. the right of individuals to have their data no longer processed and deleted when they are no longer needed for legitimate purposes. This is the case, for example, when processing is based on the person's consent and when he or she withdraws consent or when the storage period has expired;*
>
> *- complementing the rights of data subjects by ensuring 'data portability', i.e., providing the explicit right for an individual to withdraw his/her own data (e.g., his/her photos or a list of friends) from an application or service so that the withdrawn data can be transferred into another application or service, as far as technically feasible, without hindrance from the data controllers.*

The General Data Protection Regulation 2012 was voted for in the European Parliament in March 2014. At the time of writing (February 2015) it was due to be put before the Council of Ministers. To become law it has to be adopted by the Council following three-way discussions between the European Commission, the Council of Ministers and the European Parliament to develop a consolidated version of the DPR. Article 88 of the Regulation repeals

the Data Protection Directive. The Regulation will be directly implemented and will be binding on all member states.

The DPR has two stated purposes:

> *To protect the fundamental right to data protection and to guarantee the free flow of personal data between Member States*

In its introduction the Regulation refers to the Stockholm Programme which states that: *"…citizens' privacy must be preserved beyond national borders, especially by protecting personal data"* (European Council 2010). The free flow of information is seen as a pre-requisite for an open market within the EU. Lynskey (2012, pp.169–177) identified individual control over personal data as another unstated but implied purpose of the proposed DPR. The introduction to the Regulation indicates that the intention is to build trust and facilitate (online) economic activity.

Data Protection Principles

The data protection principles enshrined in the DPA have a pedigree reaching back to the OECD Guidelines (OECD 1980). These principles have been carried forward in the DPR. Table 7 shows the relationship between the data protection principles laid out in these three documents. The eight principles laid out in the DPA provide a good starting point for analysing the applicability of legislation to regulation of access to personal data.

The DPR embodies six *"Principles relating to personal data processing"* which are described in Article 5 of the Regulation. They correspond broadly to the first six of the eight principles in the Data Protection Directive (and incorporated into the UK's Data Protection Act).

The DPR makes provision for implementation across the European Union including the setting up of institutions and administrative structures to facilitate consistency in 'Chapter VII – Co-operation and Consistency'. The Regulation, which is directly binding on member states, will replace the current Directive. This provides *"a harmonised set of core rules"*.

TABLE 7 - COMPARISON OF DATA PROTECTION PRINCIPLES

OECD Guidelines	UK Data Protection Act 1998	Data Protection Regulation
Collection Limitation Principle There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject.	Principle 1 Personal data shall be processed fairly and lawfully and, in particular, shall not be processed unless— (a) at least one of the conditions in Schedule 2 is met, and (b) in the case of sensitive personal data, at least one of the conditions in Schedule 3 is also met.	(a) Personal data must be: processed lawfully, fairly and in a transparent manner in relation to the data subject
Data Quality Principle Personal data should be relevant to the purposes for which they are to be used, and, to the extent necessary for those purposes, should be accurate, complete and kept up-to-date.	Principle 3 Personal data shall be adequate, relevant and not excessive in relation to the purpose or purposes for which they are processed. Principle 4 Personal data shall be accurate and, where necessary, kept up to date.	(c) Personal data must be adequate, relevant, and limited to the minimum necessary in relation to the purposes for which they are processed; they shall only be processed if, and as long as, the purposes could not be fulfilled by processing information that does not involve personal data (d) Personal data must be accurate and kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay

OECD Guidelines	UK Data Protection Act 1998	Data Protection Regulation
Purpose Specification Principle The purposes for which personal data are collected should be specified not later than at the time of data collection and the subsequent use limited to the fulfilment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose.	Principle 2 Personal data shall be obtained only for one or more specified and lawful purposes, and shall not be further processed in any manner incompatible with that purpose or those purposes.	(b) Personal data must be collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes
Use Limitation Principle Personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with Paragraph 9 except: • a) with the consent of the data subject; or • b) by the authority of law.	*See Principle 2 above*	
Security Safeguards Principle Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorised access, destruction, use, modification or disclosure of data.	Principle 7 Appropriate technical and organisational measures shall be taken against unauthorised or unlawful processing of personal data and against accidental loss or destruction of, or damage to, personal data.	

The OECD (1980) principles include the following, which are covered in the main body of the DPA:

> *Openness Principle – There should be a general policy of openness about developments, practices and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal*

data, and the main purposes of their use, as well as the identity and usual residence of the data controller.

The following additional principles have been incorporated into the DPA:

Principle 5 – Personal data processed for any purpose or purposes shall not be kept for longer than is necessary for that purpose or those purposes.

Whereas the draft Regulation states:

Personal data must be: (e) kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed;

The DPA also upholds the rights of individuals:

Principle 6 – Personal data shall be processed in accordance with the rights of data subjects under this Act.

The Regulation's final principle contains a commitment by Data Controllers to uphold the provisions of the Regulation:

Personal data must be (f) processed under the responsibility and liability of the controller, who shall ensure and demonstrate for each processing operation the compliance with the provisions of this Regulation.

As expressed in the DPA, the applicability of each of the eight principles to SNSs is considered in turn.

PRINCIPLE 1 – FAIR AND LAWFUL PROCESSING

1 Personal data shall be processed fairly and lawfully and, in particular, shall not be processed unless—

(a) at least one of the conditions in Schedule 2 is met, and

(b) in the case of sensitive personal data, at least one of the conditions in Schedule 3 is also met.

It could be argued that the DPA does not apply to SNS providers because they are not data controllers, rather it is the users (data subjects) responsible for putting their personal data up

on SNS profiles who effectively are the data controllers. However users alone do not "*determine the purposes for which and the manner in which any personal data are, or are to be, processed*" (Data Protection Act 1998 s.1(1)). The service provider also processes personal data for the purposes of selling it to advertisers. The DPR reinforces this with its definition:

> *'controller' means the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes, conditions and means of the processing of personal data;*

At least two of the conditions in Schedule 2 (para.1, and para.6) of the DPA (referred to in Principle 1) could be argued as being met by SNSs:

> *The data subject has given his consent to the processing.*

and

> *The processing is necessary for the purposes of legitimate interests pursued by the data controller…*

In order to gain access to a service, users normally sign up to Terms and Conditions, which make provision for the service provider to utilize and exploit data provided by the user. This probably qualifies as 'consent' referred to in the DPA (sch.2, s.1). This is reinforced by the privacy policies which are analysed in Chapter 8. This not only applies (in many cases) to personal data on an individual profile, but also to information about the individual posted by other users of the service. This could include photographs of individuals automatically tagged by the system or tagged directly by individuals.

Schedule 3 of the DPA requires that "*explicit consent*" is given to the processing of sensitive personal data, which may be implied when the user signs up to the service. However given the length and complexity of many terms of service and privacy policies, it is unlikely that many individuals would have read the conditions in detail, so would not be able to provide explicit (or informed) consent. The DPD (para.33) also refers to this:

> *"Whereas data which are capable by their nature of infringing fundamental freedoms or privacy should not be processed unless the data subject gives his explicit consent"*

The DPA (sch.2, s.6) makes provision for the "*legitimate interests pursued by the data controller*" provided it does not prejudice the rights and freedoms of the data subject. As mentioned earlier, the Human Rights Act asserts "*Everyone has the right to respect for his private and family life, his home and his correspondence.*" The question then becomes: Is this right to a private life contravened by having his or her personal data passed on to third parties such as advertisers by SNS providers?

The DPR also introduces the concept of explicit consent which requires that individuals are aware of when they are giving consent for their personal data to be processed. It states that *"Silence or inactivity should therefore not constitute consent"*. (para 25). Article 7 clearly lays out the conditions for consent which stipulates that the burden of proof of consent lies with the data controller, which could be interpreted as being the SNS provider. Data subjects also have the right to withdraw their consent at any time.

PRINCIPLE 2 – EXTENSION OF PURPOSE

> *2 Personal data shall be obtained only for one or more specified and lawful purposes, and shall not be further processed in any manner incompatible with that purpose or those purposes.*

If data is gathered for the purposes of providing a social networking service to users of the service, it could be argued that selling on that data for financial gain to advertisers is in contravention of the second principle. The next question has to be whether the data that is 'sold on' to advertisers contains any personal data where a living person can be identified from the personal data or from that and other data in the possession of the data controller.

The DPA (s.10) covers the "*Right to prevent processing likely to cause damage or distress*" and could cover, for example, malicious social network entries, hounding by advertisers and spammers and stalking. Section 11 goes on to state the user's right to prevent processing for purposes of direct marketing, although the onus is on the user to first provide "*notice in writing to a data controller*" before the courts will intervene.

The data portability theme in the Article 29 Working Party statement raises some of the same issues that were discussed in the 'right to be forgotten' proposal in the EU ePrivacy Directive 2009. Transfer of personal data from one service to another implies that it is removed from the source after it has been copied to the destination. For the reasons described above, there is no absolute guarantee that the data can ever be totally removed from the source. If the personal data has ever been transmitted via a satellite communications relay, signal

leakage means that extra-terrestrial measures (and faster than light travel) would be required to eliminate all copies of that personal data. In other words it is impossible to guarantee the elimination of all data once it has been stored or transmitted (Vsauce 2013).

Opinion 2/2010 of the European Commission on online behavioural advertising focuses specifically on the issue of cookies or other tracking devices placed in browsers to follow the behaviour of users online (Justice 2010). The Opinion highlights the issue of informed consent and recommends an 'opt-in' principle should be implemented. It also highlights the fact that ad network providers and website publishers effectively become data controllers and have some data protection responsibilities. It is of the view that "*creation of very detailed user profiles* [...] *in most cases, will be deemed personal data*". It states that:

> "*Ad network providers should: i) limit in time the scope of the consent: ii) offer the possibility to revoke it easily and iii), create visible tools to be displayed where the monitoring takes place.*"

This reinforces the "*right to refuse*" mentioned in Testimonial 66 of the ePrivacy Directive 2009 when discussing cookies.

PRINCIPLE 3 – ADEQUATE AND RELEVANT DATA

> *3 Personal data shall be adequate, relevant and not excessive in relation to the purpose or purposes for which they are processed.*

Principle 3 is problematical because personal data is not clearly defined in the context of SNS providers. Data put up on personal profiles in many social networks falls within the definition of sensitive personal data DPA (s.2) including:

- Racial origin
- Political opinions
- Religious beliefs
- Physical or mental health
- Sexuality

However a lot of the information put up is of a fanciful nature – and may not be accurate or necessarily true. In these cases one might legitimately question whether it qualifies as personal data. For example the Section 32 exemption for literary work might apply. If this is the case, copyright law may also apply. It also introduces the question: 'Is the data subject under an obligation to provide accurate or truthful data in these services?' Some of the

terms of service require accurate personal data, although there is often no easy way of policing this.

Data minimisation is incorporated into the six principles in Article 5 of the General Data Protection Regulation 2012, which states:

> *Personal data must be: adequate, relevant and limited to the minimum necessary in relation to the purposes for which they are processed*

PRINCIPLE 4 – ACCURATE AND UP-TO-DATE DATA

> *4 Personal data shall be accurate and, where necessary, kept up to date.*

Section 14 of the DPA on "*Rectification, blocking, erasure and destruction*" presents difficulties in that a court may not be able to "*order the data controller to rectify, block, erase or destroy those data and any other personal data ... which appears…to be based on the inaccurate data*." For instance, if the data has been entered by the data subject and is (at least partly) based on fantasy, it would be difficult to apply this provision.

PRINCIPLE 5 – RETENTION

> *5 Personal data processed for any purpose or purposes shall not be kept for longer than is necessary for that purpose or those purposes.*

Although it is clear that personal data should not be kept longer than necessary, some SNSs make provision in their privacy policies for ongoing preservation of personal data for a variety of reasons including operational, backup and legal compliance purposes. These are discussed in more detail in Chapter 8.

The fifth principle in the proposed DPR (art.5) states that data must be kept *"for no longer than is necessary for the purposes for which the personal data are processed"*, although it does make provision for data *"processed solely for historical, statistical or scientific research purposes"*. This implies that if an individual no longer wishes personal data to be kept on the social network, they can have it removed.

PRINCIPLE 6 – RIGHTS OF DATA SUBJECTS

> *6 Personal data shall be processed in accordance with the rights of data subjects under this Act.*

Principle 6 of the DPA reinforces the idea that the rights of users should be taken into account in the processing of personal data. This is also addressed in Article 1 of the proposed Regulation:

> *This Regulation protects the fundamental rights and freedoms of natural persons, and in particular their right to the protection of personal data.*

PRINCIPLE 7 – UNAUTHORISED ACCESS

> *7 Appropriate technical and organisational measures shall be taken against unauthorised or unlawful processing of personal data and against accidental loss or destruction of, or damage to, personal data.*

Recent European directives on electronic communications have modified the scope of the DPA to regulate electronic monitoring and advertising. The ePrivacy Directive 2009 makes specific provision for regulating use of cookies and offering users the right to refuse them and a user-friendly manner (Recital 66). The directive also refers to "*unsolicited commercial communications (spam)*" to allow internet service providers to initiate legal proceedings against spammers. This has a direct impact on the majority of SNS providers.

The DPR covers enforcement comprehensively in Chapter VIII – Remedies, Liability and Sanctions.

PRINCIPLE 8 – TRANSFER OF DATA BEYOND THE EEA

> *8 Personal data shall not be transferred to a country or territory outside the European Economic Area unless that country or territory ensures an adequate level of protection for the rights and freedoms of data subjects in relation to the processing of personal data.*

There is a clear correlation between the OECD principles and those incorporated into the DPA and carried forward in the proposed Regulation. However the DPA goes further in developing the concept of trans-border data flow and retention of personal data which are covered in Principle 8.

Principle 8 is intended to prevent transfer of personal data outside the EEA unless there is "*an adequate level of protection ... of personal data*". However if the social network provider is based outside the EU, the question of transfer of personal data may be difficult to enforce.

Section 5 deals with application of the DPA and states that:

> *"...this Act applies to a data controller in respect of any data only if –*
>
> *The data controller is established in the United Kingdom and the data are processed in the context of that establishment, or*
>
> *the data controller is established neither in the United Kingdom nor in any other EEA State but uses equipment in the United Kingdom for processing the data ..."*

As many social network providers are not based in the UK and deliver their services via the internet, some take the view that the Act does not apply to their services. In other words if the Data Controller is not in the UK or an EEA country, the Act does not apply to them. However paragraph 3 (d) states that the Act applies to:

> *"any person who* [...] *maintains in the United Kingdom –*
>
> *An office, branch or agency through which he carries on any activity, or a regular practice"*

This raises the question of what constitutes "*an office, branch or agency*" or a "*regular practice*" and whether provision of an internet service available within the UK falls within this definition.

Some of the major SNS providers such as Facebook, LinkedIn and Google are based in the United States, although many also have offices in the European Union. US operators are covered by the U.S.-EU Safe Harbor framework, which allows transfer of personal data to US-based organisations. However its scope and mode of operation is significantly different from the workings of the DPD. In particular there is a concern about its self-regulatory nature, the lack of active enforcement of its provisions, and the lack of any compulsion for independent certification of compliance. The TRUSTe service is one of the main independent certifying bodies.

The principle of transfer of data to countries or territories outside the European Union is covered by the DRP in 'Chapter V – Transfer of Personal Data to Third Countries or International Organisations'. Article 3 clearly states that the Regulation applies to the personal data of EU residents where services are offered in the Union or their behaviour is monitored. Chapter V deals with issues of territoriality, including a requirement that the third country or international organisation provides an adequate level of protection of personal data. A list of complying countries will be published in the *Official Journal of the European Union* when the Regulation comes into effect.

Issues Arising from the General Data Protection Regulation

The Right To Be Forgotten

There is the wider issue of policing access to personal data and the fact that it is impossible to truly delete anything from the internet. Once information is in the public domain on the internet or anywhere else, it would seem an insurmountable problem to identify every instance of that data and to require its deletion. In addition there are back-ups to websites – required for operational purposes to which it would be impractical to retroactively apply a deletion and digital archiving projects where the internet or parts of the internet are being archived for study and as a primary resource for future researchers.

The Opinion from the European Commission (2010) suggests that personal data is viewed in some respects like a physical entity that has a specific location and can consequently be transferred or destroyed (erased). This belies the enduring nature of data on the internet and the nature of knowledge – that once discovered, it is impossible to deliberately 'undiscover' information. This would mean that any legislation requiring data controllers to enact the 'right to be forgotten' will be impossible, particularly in the context of social networks and other environments where personal data may be widely distributed.

The 'right to be forgotten' may work better in a controlled environment such as within a company dealing with employee or customer information. Even then, with most data being held electronically, it would be very difficult to ensure that all copies, back-ups and versions of personal data had been removed. The service providers will almost certainly have their own back-up and archiving procedures in place, so that even if data is removed at a user's request it will still persist (although it may not be available online and would effectively be inaccessible). This makes it very difficult to guarantee imposition of the terms of the 'right to be forgotten'.

The 'right to be forgotten' is enshrined in the DPR. This has become a major point of discussion following the European Court of Justice Ruling in May 2014 upholding the Spanish Court's requirement that Google Spain remove links to a newspaper article about a named individual who objected to being associated with out of date information about him (Haynes 2014a; Court of Justice of the European Union 2014). Para 53 states *"Any person should have the right to have personal data concerning them rectified and a 'right to be forgotten'"*. It does however make a number of exceptions such as retention for *"historical, statistical and scientific research purposes"*. Article 17 of the DPR specifically states:

> *The data subject shall have the right to obtain from the controller the erasure of personal data relating to them and to the abstention from further dissemination of such data*

The 'right to be forgotten' only applies to living individuals and this can lead to problems getting the profiles of deceased users removed from SNSs.

DATA PORTABILITY

One of the most interesting provisions from the perspective of SNSs is the right of data portability. In the introductory text of the DPR (para.55) social networks are specifically mentioned as an example of data portability. This suggests that they are a particular target and will be actively monitored for compliance by the authorities when the Regulation comes into force. Article 18 specifies:

> *Where the data subject has provided the personal data and the processing is based on consent or on a contract, the data subject shall have the right to transmit those personal data and any other information provided by the data subject and retained by an automated processing system, into another one, in an electronic format which is commonly used, without hindrance from the controller from whom the personal data are withdrawn.*

ENFORCEMENT

The Information Commissioner has powers under the DPA and related statutory instruments to enforce the principles of data protection in the UK. The role of the Information Commissioner's Office (2014) is "*to uphold information rights in the public interest, promoting openness by public bodies and data privacy for individuals*".

The Information Commissioner has the power to impose a monetary penalty on a data controller if he or she deliberately contravenes the data protection principles. It could be argued that selling personal data from profiles put up by users on SNSs would be such a contravention.

The UK's Data Protection (Monetary Penalties) (Maximum Penalty and Notices) Regulations (SI 2010/31) raised the maximum fine that the Information Commissioner could impose to £500,000 for breaches of the DPA. Under the Regulation this will increase to a fine of €1 million or 2% of the annual worldwide turnover of an enterprise for the most serious infractions.

The ICO has also been active in targeting serious breaches for public sanction by 'naming and shaming' offenders through news releases and public notices on its website: www.ico.gov.uk/enforcement.

Under Article 31 of the proposed Regulation, data controllers will be required to notify the supervisory authority about a personal data breach. Article 32 specifies that data subject directly affected by a data breach should also be notified. 'Chapter VIII – Remedies, Liability and Sanctions' gives data subjects the right to lodge complaints and the right to a judicial remedy.

ADMINISTRATIVE BURDEN

Although there would no longer be a requirement to keep a data protection register, the Regulation is prescriptive in the way in which data controllers in companies and public sector organisations operate. They are expected to respond to data subject's need to be informed. Under Article 15 they are required to provide information on *"the storage periods, and of the rights to rectification and to erasure and to lodge a complaint"*. Chapter IV of the regulation goes on to set out the responsibilities of the data controller (Articles 22), the processor (Article 26) and the data protection officer (Article 5, 35-37), which should be appointed by public sector bodies and large enterprises.

Although it could be argued that individual members of an online SNS have a degree of control over the purposes of processing their own personal data, it is difficult to argue that they determine the "*conditions or means of the processing"*. This is a role for the SNS providers. This then introduces the next problem of who the provider actually is. Where a company is headquartered outside the EU but provides services to EU residents, there is some responsibility under the Regulation to protect individuals' personal data. Where there is a

subsidiary in the European Union, they may be treated as controllers of the personal data, even if the data is hosted on a server beyond the EU's physical boundaries.

Concerns about administrative burdens on micro, small and medium-sized enterprises are addressed in Article 22 which talks about proportionality of *"measures to ensure and be able to demonstrate that the processing of personal data is performed in compliance with this Regulation."*

EXEMPTIONS

Part IV of the DPA makes specific exemptions, two of which were highlighted in a telephone interview with an advisor from the Information Commissioner's Office in 2011. They are: Section 32 'Journalism, literature and art' and Section 36 'Domestic purposes'.

The section 32 exemption relates to "*processing ...undertaken with a view to the publication by any person of any journalistic, literary or artistic material*". It goes on to define 'publish' as: "*make available to the public or any section of the public*." Clearly putting up a personal profile on a social network would qualify under this definition, if it becomes visible either to the community of subscribers to that service, or to a wider internet audience through search engines such as Google. Some social networks allow users to restrict the visibility of their profiles to a group selected by them, which could be argued as falling outside the exemption.

The second aspect of this is what constitutes "*journalistic, literary or artistic material*". For instance, is it legitimate to use personal details on an SNS profile for an article published in a newspaper or magazine? This may be an issue if the consent of the data subject has not been obtained. It could also be seen as an additional purpose for which the data was not initially gathered.

Section 36 refers to "*Personal data processed by an individual only for the purposes of that individual's personal, family or household affairs (including recreational purposes)*". Here the question is whether the data is solely processed by the individual user of the service (data subject), or whether it is also processed by the service provider and therefore whether it is exempt from the data protection principles. Article 3 of the Directive states:

> *This Directive shall not apply to the processing of personal data:* […] *-by a natural person in the course of a purely personal or household activity.*

Bond (2010) raises the issue of data ownership and, specifically, who the data controller is. He suggests that social network service providers as well as individual users of SNSs have obligations as data controllers. He goes on to say:

> *The fact that there is a business driver to the use of social media means* […] *that the household exemption does not apply and so the question arises as to whether or not the privacy policy notification and other practices of an organisation that is using a service such as LinkedIn for business development purposes are sufficient.*

Section 55 of the Act covers "*unlawful obtaining etc. of personal data*" and states:

> *(1) A person must not knowingly or recklessly, without the consent of the data controller-*
>
> *(a) obtain or disclose personal data or the information contained in personal data, or*
>
> *(b) procure the disclosure to another person of the information contained in personal data.*

However if the data controller is also the social network service provider, he or she is giving consent to pass on personal data to advertisers for the purposes of direct marketing.

Many SNS providers include exceptions to their privacy policies (see Chapter 8) which to some extent accord with the proposed Regulation. Article 21 of the DPR allows for restriction of the scope of specified rights and obligations to safeguard security or *"the prevention, investigation, detection and prosecution of criminal offences"*. The other exceptions are not usually cited in privacy policies. The other major exemption is the processing for historical, statistical and scientific research purposes so long as the processing of the data *"does not permit or not any longer permit the identification of the data subject"*. This is an important consideration for online behavioural advertising, where personal data is aggregated for the purposes of identifying target groups for advertisements. As this is a major source of income for SNSs, this becomes a crucial issue for the continued viability of the dominant economic model for social media.

RISK MANAGEMENT

In the introductory section of the DPR the role of risk management in protecting individuals is acknowledged by setting out some broad categories of risk and outlining risk management framework in very general terms. Section 2, Chapter IV of the DPR deals with data security and specifies that *"the processor shall implement appropriate technical and organisational measures to ensure a level of security appropriate to the risks…"*. It identifies specific risks that are discussed in Chapter 5 and requires consideration of the consequences and ways of mitigating the adverse effects of a personal data breach.

PRIVACY BY DESIGN AND USER EDUCATION AS MEANS OF REGULATION

Lessig (2006) suggests that legislation alone is insufficient to regulate the internet. This principle also applies to the specific area of protecting personal data on SNSs and is acknowledged in Article23 and Article 52 of the DPR. Article 23 deals with 'Data protection by design and by default' and requires the data controller to *"implement appropriate technical and organisation measure and procedures …* [to] *ensure the protection of the rights of the data subject."* The DPR (art.52, para.2) goes on to say:

> *Each supervisory authority shall promote the awareness of the public on risks, rules, safeguards and rights in relation to the processing of personal data.*

This means that national regulators such as the Information Commissioner's Office in the UK have a responsibility to educate the public about risks and online safety. As well as monitoring compliance and taking action where there are breaches of the Act, the Information Commissioner provides advice and guidance. The ICO also has an educative role offering guidelines for young people on protecting their personal data (Information Commissioner's Office 2009).

The ICO publishes guidelines for organisations about specific data protection issues as well as guidelines targeted at different sectors, such as: business, finance, charities, health, education, local authorities, marketing, and MPs and political parties.

The concept of 'privacy by design' has been around for a long time, and is implied in the DPD which refers to: "*appropriate technical and organizational measures* […] *at the time of the design of the processing system* […] *in order to maintain security and thereby to prevent any unauthorized processing*" (para.46). The ICO has issued guidelines, which do not have force

of law, but which are intended to represent good practice for service providers and systems developers. Privacy should be one of the primary considerations when a new information system that handles personal data is being developed. A review funded by the Information Commissioner concluded that privacy by design was increasingly being recognised by data protection authorities as a way of addressing growing concerns about privacy in information systems. The approach seems to work best "*When they are part of a system of incentives, sanctions and review*" (Information Commissioner's Office 2008).

CONCLUSION

A review of the legislation in the UK as it applies to personal data on SNSs suggests that the Data Protection Act 1998 is the single most relevant piece of legislation. This in turn is based on the EU's Data Protection Directive (95/46/EC) and associated legislation. It is also dependent on the Human Rights Act 1998 which requires the right to privacy. The DPA is the UK implementation of the Directive which will be repealed by the General Data Protection Regulation 2012 when it comes into force. The Regulation embodies several new principles that are likely to affect SNS operations in the EU:

- Data minimisation
- Data portability
- Right to be forgotten

There is a problem of interpretation of the DPA and whether it applies to social media providers. Objections to application of the Act focus on the exemptions described in Article 32 *"publication…of any journalistic, literary or artistic material"* and section 36 which refers to data processed *"only for the purposes of that individual's personal, family or household affairs"*. This research argues that the exemptions do not apply to SNSs, because it is not the intent of most users to publish their personal data. Nor is the data being used solely for personal purposes, as clearly many social media providers are using the data for advertising and other purposes.

Attempts to define users as 'data controllers' imply that users have complete control over the processing of the personal data. This is not the case, because the SNS providers (who could be argued to be the true data controllers) are processing and using personal data in ways not envisaged by users. Although users have access to privacy policies and have to sign up to the terms of service, it is difficult to argue that this qualifies as informed consent, as they are rarely read through and are often lengthy and difficult to absorb in one go.

Another objection to the application of the DPA to social media providers based in the United States, is that they are in theory covered by the U.S.-EU Safe Harbor framework. This self-regulatory scheme is discussed in Chapter 8.

Future developments in legislation are intended to cover novel uses of personal data such as use of internet cookies and sending unsolicited e-mails to individuals. While these cover some of the areas of potential misuse of personal data, the provisions as currently drafted are not realistic in their expectation that personal data can be deleted or 'forgotten'.

This leads to the conclusion that while there are important principles in the DPA, it is insufficient as a tool to effectively regulate access to personal data on SNSs. At the very least it would have to be used in conjunction with other regulatory modes.

CHAPTER 8 – SELF-REGULATION: PRIVACY POLICIES

INTRODUCTION

Brown and Marsden (2013, p.12) talk about self-regulation in terms of norms and code: *"Self-regulation is defined broadly as a rule of the formation of norms: it exercises a function that shapes or controls the behaviour of actors in that environment, which may include software code."* However the privacy policies published by SNS providers represent a self-regulatory approach to protecting users' interests and can be seen as an expression of both the code and the norms that shaped that code. They are also subject to legislative regulation as may be seen by the 2012 findings by CNIL, the French regulators, on Google's consolidated privacy policy for all its public services (CNIL 2012; European Privacy Authorities 2012).

Boyd and Hargittai (2010) report on concerns about privacy settings on Facebook and suggest that contrary to many assumptions, users of Facebook are concerned about the risks associated with personal information on SNSs. This reflects more general concern about what personal data is held by SNS providers and the way in which it is being used. Access to personal data on social networks is regulated in a number of ways. This chapter looks at one of these: self-regulation and its expression in the privacy policies of the main SNS providers. It goes on to consider the relationship between privacy policies and legislation.

Jimenez (2009) identified the need for good privacy policies and some way of helping users to distinguish between the different levels of privacy protection available on social media, although he recognised the need for independent watchdogs to ensure that the self-regulation works.

Even where privacy policies exist, there is no guarantee that users read them. For instance, a survey of students at one college in the US found that 47% of users had not read Facebook's privacy policy at all (Stutzman et al. 2011, p.592) and another college-based study found that 41% had not read the privacy policy (Butler 2011, p.49). This is in line with another study, which found that 45% of users had not read the privacy policies (Levin & Abril 2009). The college study went on to find that reading privacy policies tended to reduce the levels of disclosure, but that this was more than offset by personalising privacy settings which increased disclosure. Although the study was limited in extent and generalizability it does provide an interesting avenue for further investigation (Stutzman et al. 2011, pp.596–597).

This investigation is based on a selection of privacy policies downloaded into NVivo10 in July 2014, and converted to Word files for content analysis. This follows an earlier study of policies in October 2011. The content was coded using a combination of text searches and manual coding. Some of the codes were identified directly from the text of the policies and others were applied based on the overall questions that this research set out to answer. This provided a basis for collating policy extracts for comparison and analysis. The review of privacy policies set out to answer the following questions:

- To what extent do privacy policies regulate access to personal data?
- What protections do privacy polices offer to UK users of SNSs?
- How do privacy polices relate to other forms of regulation?
- Is there a way of enforcing privacy policies?

The privacy policies of the following SNSs were selected on the basis that they were:

- Online social networking services (SNSs) with personal profiles and an ability to connect with other users
- Not primarily media- or file-sharing sites that do not require membership (e.g. YouTube)
- They were available in the English language and used by UK-based members
- They were in the top ranked global websites as defined by the Alexa page rankings (Alexa 2014)

The privacy policies of eleven widely-used, global SNSs available to English-speakers were reviewed in July 2014. The rankings were initially identified by Wikipedia (2014) entry which referred to the Alexa page rankings. Table 8 contains updated Alexa rankings from 2014. The actual ranking may not be too critical – as long as the major providers have been picked up in this way. The numbers of registered users were mostly taken from the Wikipedia page, although where absent, the SNS website was consulted (this is indicated by an asterisk). The method of ranking is more significant when smaller, more specialist social network services are involved.

TABLE 8 - SIZES OF SNSS

Site	No. of registered users	Alexa ranking 2014	Alexa ranking 2011
Badoo	219 million*	188	117
Facebook	1.28 billion	2	2
Google	1.6 billion	1	
hi5	80 million	1,379	
Instagram	150 million	30	
LinkedIn	255 million	12	13
Myspace	30 million	1,080	131
Ning	2 million communities*	622	273
Snapchat[3]	70 million MAUs[4]	16,571	
Twitter	94 million	7	9
WhatsApp	500 million MAUs*	2,658	

PRIVACY POLICIES

The privacy policies were downloaded in late July 2014. Two of the eleven policies were more than a year old, five were between 6 and 12 months and four were less than 6 months old. This suggests that privacy policies are still being actively updated and that they have not stabilised. In some instances there were links back to previous versions of the privacy policy.

Changes to privacy policy are a perennial problem and the cause of misunderstandings between users and providers (Kuzma 2011; Butler 2011). There is often a discrepancy between the privacy protections that users believe they have and the actual protections stated in the privacy policy. One possible explanation may be the rate at which the privacy policy changes, making it difficult for users to keep up (Butler 2011; McKeon 2010; Wilson et al. 2012).

Most SNSs made some commitment to notify users of changes to the privacy policy, in some cases if the changes resulted in reduced privacy. The means of notification was not always specified and where it was, it was usually limited to a commitment to post the change on the privacy policy page. Some, such as Myspace did undertake to notify users directly:

[3] (Edwards 2014)

[4] Monthly Active Users

> *…if we make material changes to this Policy that expand our rights to use your Personal Information, we will notify you either through a Myspace message, e-mail, and/or a prominent notice at the Myspace Website.*

In the majority of cases, continued use after notification implied acceptance of the new privacy policies. For example, LinkedIn's policy says: *"Using LinkedIn after a notice of changes has been communicated to you or published on our Services shall constitute consent to the changed terms or practices."* Only Facebook offered the opportunity to comment on proposed changes: *"Unless we make a change for legal or administrative reasons, or to correct an inaccurate statement, we will give you seven (7) days to provide us with comments on the change."*

Change of ownership may also affect privacy policies and most SNSs were at pains to emphasise they had the right to transfer personal data to the new owners. For example:

> *In the event that Badoo or any of its affiliates undergoes a business transition or change of ownership, such as a merger, acquisition by another company, re-organisation, or sale of all or a portion of its assets, or in the event of insolvency or administration, we may be required to disclose your personal information.*

> *Myspace may disclose User information in the event we file for bankruptcy, or in connection with certain types of business transactions that result in corporate change. For example, if Myspace is or may be (or substantially all of its assets are or may be) acquired by a third party pursuant to an acquisition, merger, sale, reorganization, liquidation, or similar business transfer, we may disclose and transfer User information to the parties involved in the transaction (e.g., in connection with due diligence efforts or as a transferred asset).*

And some give a commitment to maintaining the protections offered by the privacy policy under the new ownership:

> *If the ownership of our business* [Facebook] *changes, we may transfer your information to the new owner so they can continue to operate the service. But they will still have to honor the commitments we have made in this Data Use Policy.*

> *If Google is involved in a merger, acquisition or asset sale, we will continue to ensure the confidentiality of any personal information and give affected users notice before personal information is transferred or becomes subject to a different privacy policy.*

The scope of the policies sometimes indicated that they were in conjunction with the general terms and conditions of use: *"This Privacy Policy is incorporated into and is subject to the Ning Terms of Service."* Many also stated that usage of the service implied acceptance of the terms of the privacy policy: *"When using any of our Services you consent to the collection, transfer, manipulation, storage, disclosure and other uses of your information as described in this Privacy Policy."*

SNSs were also keen to show the exclusions:

> *Our Privacy Policy does not apply to services offered by other companies or individuals, including products or sites that may be displayed to you in search results, sites that may include Google services, or other sites linked from our services. Our Privacy Policy does not cover the information practices of other companies and organizations who advertise our services, and who may use cookies, pixel tags and other technologies to serve and offer relevant ads.*

> *The activities of third parties or IMH affiliates that provide services (including advertising services) to Myspace, or that you link to from us, or that may collect information about you themselves or that we may share your information with, including our affiliates such as Specific Media LLC ("Specific Media"), are governed by the respective privacy policies of those third parties or IMH affiliates.*

> *This Privacy Policy does not apply to the practices of companies that WhatsApp does not own or control, or to individuals whom WhatsApp does not employ or manage, including any of the third parties to which WhatsApp may disclose user information as set forth in this Privacy Policy.*

Several services including Google and hi5 included definitions of some of the terms appearing in the privacy policies or included links to sites with fuller explanations of terms and concepts.

Most provided contact details either as a link or with an e-mail address and often also a postal address.

PERSONAL INFORMATION

The SNSs collect a variety of personal information and this can include transactional data and content associated with an individual profile. Most of the privacy policies give examples of the types of data that they collect, but they are not necessarily exhaustive. Nevertheless it is instructive to see where the similarities lie and to detect any patterns in the types of data gathered. Table 9 summarises the main data types identified. There is some discussion about what constitutes personal data – discussed earlier, with some SNSs indicating that aggregated data and some transactions data does not (in their view) count as personal data. Not all the information requested at registration (e.g. Google request for a photo, LinkedIn request for gender information) was mandatory.

TABLE 9 - MINIMUM REGISTRATION DATA GATHERED

	Badoo	Facebook	Google	hi5	Instgram	LinkedIn	Myspace	Ning	Snapchat	Twitter	Whatsapp
Name	✓	✓	✓	✓		✓	✓	✓		✓	
E-mail	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Gender	✓			✓		✓	✓				
DoB/Birthday	✓	✓		✓			✓				
Location (GPS)	✓			✓							
Address				✓		✓	✓				
Financial (payment card)	✓		✓					✓			
IP Address								✓			
Mobile/Telephone Number		✓	✓			✓	✓	✓			
Photo			✓								
Username					✓		✓			✓	
Password					✓	✓		✓	✓	✓	

A lot of tracking and logging information is also recorded for a variety of purposes. Table 10 shows the types of information routinely gathered by SNSs. Online behavioural advertising is most closely associated with the tracking data whereas the logging data ("L" in the table) is often used as part of service provision. (Key T = Tracking and cookies, D = Device information, L = Logging/Transactional data)

TABLE 10 - TRACKING DATA GATHERED BY SNSS

	Badoo	Facebook	Google	hi5	Instgram	LinkedIn	Myspace	Ning	Twitter
DEVICE									
IP Address or Device ID	T	D/T	D/L		D	D	D	L	L
Computer		D	D			D	D		
Mobile Phone		D	D						L
Operating System		D/T	D			D	D		L
Cookies		T	L	L/T		T	T	L	L
ACTIVITIES									
Messaging and Communications	T	L	L	L					L
Add-ons					T				
Friends' Details	T	L		L					
Networking Groups									L
Search / Online activity		L	L	T		L	L		L
Page Visits		D/T		T	T			L	L
PERSONAL IDENTIFIERS									
User ID and Social log-in		T		L		L			

ANONYMISED DATA

Many SNSs make the distinction between personal data and that which has been anonymised in some way (for instance by aggregating it and stripping out personal identifiers). This particularly is applied to tracking information and device identifiers that some providers consider distinct from personal data. The ICO provides extensive treatment of the risks of re-identification (de-anonymisation) from aggregated or otherwise anonymised personal data (Information Commissioner's Office 2012).

Aggregated data consists of system and behavioural data gathered by the SNS providers, such as IP address, browser used, sites visited, options selected, as well as general categories (sometimes self-defined) and characteristics selected by advertisers such as: age, gender, occupation, location, and interests. Privacy policies describe anonymised or aggregated data in terms of what is done with it:

> *Badoo discloses aggregated non-personal data for marketing and promotional purposes. That means we do not disclose any information that could be used to identify you.*
>
> *We* [Google] *may share aggregated, non-personally identifiable information publicly and with our partners – like publishers, advertisers or connected sites.*
>
> *We may also produce reports for advertisers that aggregate, in an anonymous manner, information about your and other Site users' activity on hi5.*
>
> *LinkedIn may provide reports containing aggregated impression information to companies hosting LinkedIn plugins and similar technologies to help them measure LinkedIn-generated traffic to their websites.*
>
> *Please note that this Policy does not limit Myspace's use or sharing of aggregated, de-identified or other anonymized information that does not constitute Personal Information*
>
> *We* [Snapchat] *may also share aggregated or de-identified information, which cannot reasonably be used to identify you.*

Some privacy policies did define anonymised data in the following terms:

> *We* [Instagram] *may remove parts of data that can identify you and share anonymized data with other parties. We may also combine your information with other information in a way that it is no longer associated with you and share that aggregated information.*
>
> *The term "Demographic Information" refers to information about groups or particular types of individuals that access the Myspace Services. Demographic Information does not directly identify you as a specific individual, and would*

include information such as your age, gender, marital status, income level, and zip code.

The information that we collect with these automated methods may include, for example, your IP address, Ning cookie information, a unique device or user ID, browser type, system type, the content and pages that you access on the Ning Platform, and the "referring URL" (i.e., the page from which you navigated to the Ning Platform)

We may share or disclose your non-private, aggregated or otherwise non-personal information, such as your public user profile information, public Tweets, the people you follow or that follow you, or the number of users who clicked on a particular link (even if only one did).

…non-personally-identifiable information (such as anonymous user usage data, cookies, IP addresses, browser type, clickstream data, etc.) [WhatsApp]

TRACKING TECHNOLOGIES

Tracking technology was mentioned in all of the privacy policies. Increased awareness of this as an issue may be due in part to the ePrivacy Directive 2009, which was implemented in 2012. Table 11 shows which tracking technologies are mentioned in the privacy policies. However the fact that a technology was not mentioned does not mean that the technology is not used.

TABLE 11 - TRACKING TECHNOLOGIES USED

	Badoo	Facebook	Google	hi5	Instgram	LinkedIn	Myspace	Ning	Snapchat	Twitter	WhatsApp
Cookies	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Persistent Cookies	✓	✓	✓	✓	✓		✓				✓
Flash Cookies							✓				
Web Beacons					✓	✓	✓	✓	✓		
Log File			✓	✓	✓	✓	✓				
Device identifiers			✓			✓	✓	✓		✓	✓
Third Party Cookies	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Blocking Cookies mentioned	✓	✓	✓	✓		✓	✓		✓	✓	✓

Although cookies were mentioned in all the privacy policies, the level of detail provided varied enormously. They were variously described as session cookies, authentication cookies, ID cookies, web cookies and browser cookies. They are used to authenticate users, ensure session continuity in case of interruption, and enable some of the interactive features of the site. These session cookies do not persist beyond the end of the session.

Persistent cookies were acknowledged in seven of the policies reviewed. These are also known as 'tracking cookies' or 'local storage devices'. These cookies remain on the user's browser and enable preferences to be remembered between sessions. They may also be placed on a user's browser by a third party (such as an advertising partner) in order to keep track of web pages a user is visiting. Flash cookies were a particular type of persistent cookie that can re-establish itself on a user's browser after it has been deleted. They have become less popular as they have been perceived as being a nuisance by many users. Only the Myspace privacy policy mentioned them.

Web beacons, variously known as pixel tags, tracking pixels and clear gifs, were mentioned in five privacy policies. They consist of a small (invisible) image file that can be used for tracking e-mail responses.

Log files and clickstream may be used by SNS providers or third parties for advertising as well as analysis of site usage. Five of the eleven privacy policies included statements about this type of tracking technology. Six policies (including three of the above) mentioned device identifiers such as IP addresses or Unique Application Numbers.

All the privacy policies except WhatsApp acknowledged the use of tracking information by third parties. In most cases they warned that the use of data gathered by third parties was beyond the control of the SNS service and that users were advised to consult the privacy policies of the third parties. The third parties might be companies within the same group as the SNS provider, advertisement networks, or advertisers themselves.

All but two (Instagram and Ning) mentioned that cookies could be blocked or disabled. Many were keen to point out that by doing so users may be depriving themselves of the full functionality of the SNS or access to all its features.

CONTESTS, SURVEYS AND POLLS

Contests, surveys and polls are also used as ways of enticing users to share more personal information.

> *Surveys may ask for your contact, demographic or unique identifying information to increase the value of the results. Contests may require your contact information or other demographic or personal information to determine eligibility.* [hi5]

> *Polls and Surveys may be conducted by LinkedIn, Members, or third parties. Some third parties may target advertisements to you on the results page based on your answers in the poll.*

> *We may offer sweepstakes, contests, and other promotions (collectively, "Promotions") through the Myspace Services that may require registration.* […] *If you choose to enter a Promotion, Personal Information may be disclosed to third parties or the public in connection with the administration of such Promotion…*

THIRD PARTIES

Third parties play an important role in the delivery of SNSs to users and include advertising networks that provide the majority of income for many SNS providers. This is openly acknowledged and bodies such as the Internet Advertising Bureau, UK (IABUK) provide consumer-oriented information that enables users to make more informed choices relating to tracking technology.

SHARING PERSONAL DATA

Most policies acknowledged that SNS providers share data with third parties, although most also pointed out that this data was usually aggregated to prevent identification of individual users:

> *We* [Facebook] *only provide data to our advertising partners or customers after we have removed your name and any other personally identifying information from it, or have combined it with other people's data in a way that it no longer personally identifies you.*

> *We* [Google] *may share aggregated, non-personally identifiable information publicly and with our partners – like publishers, advertisers or connected sites.*

> *From time to time, we* [hi5] *may share your email address [and/or other personal information] with third parties for marketing purposes.*
>
> *We* [LinkedIn] *do not share personally identifiable information with any third-party advertiser or ad network*
>
> *Third parties – including linked services (e.g., Twitter), advertisers, and advertising service providers may also collect certain information about you in connection with your use of the Myspace Services (e.g., Usage Information).*
>
> *…we* [Myspace] *may share your Personal Information, Demographic Information, Usage Information, Device Identifiers, and any other information that we receive or collect with third parties (including IMH affiliates and our third-party advertising partners) in order to enable customization of the advertising…*
>
> *Network Creators and Third Party Application Developers might use the Ning Platform to collect, use, and retain all information, including Personal Information and Network Data, that is disclosed and uploaded by Members in connection with their use of Networks*
>
> *We may share or disclose your information at your direction, such as when you authorize a third-party web client or application to access your Twitter account.*

Badoo does not sell personal data to third parties. However, "*Badoo discloses aggregated non-personal data for marketing and promotional purposes"*.

Some such as Google and LinkedIn indicate the circumstances under which they provide personal data to third parties. For example:

> *We protect your personal information and will only provide it to third parties: (1) with your consent; (2) where it is necessary to carry out your instructions; (3) as reasonably necessary in order to provide our features and functionality to you; (4) when we reasonably believe it is required by law, subpoena or other legal process; or (5) as necessary to enforce our User Agreement or protect the rights, property, or safety of LinkedIn, its Members and Visitors, and the public.*

Several other privacy policies also refer to the right to make personal data available to law enforcement agencies or as part of legal obligations or court proceedings. Although not specified in any of the policies viewed, this could include a requirement to make personal data available to security services and agencies in the United States under the Homeland Security Act 2002. This issue has been discussed extensively by the Electronic Frontier Foundation on its website (www.eff.org).

PROCESSORS OR SERVICE PROVIDERS

Most of the privacy policies also allowed for sharing of data for processing on their behalf by external agencies:

> *We* [Google] *provide personal information to our affiliates or other trusted businesses or persons to process it for us, based on our instructions and in compliance with our Privacy Policy.*
>
> *We* [Instagram] *also may share your information as well as information from tools like cookies, log files, and device identifiers and location data, with third-party organizations that help us provide the Service to you ("Service Providers").*
>
> *We may employ third party companies and individuals to facilitate our Services (e.g. maintenance, analysis, audit, marketing and development). These third parties have limited access to your information only to perform these tasks on our behalf and are obligated to LinkedIn not to disclose or use it for other purposes.*
>
> *We* [Myspace] *may make your information available to our agents and service providers so that they can provide requested services on our behalf.*
>
> *We may provide Personal Information to third parties for their use in performing internal business functions both for the Ning Platform and Web site (e.g., payment processing, maintenance, security, data analysis, email transmission, or data hosting) on our behalf.*
>
> *We* [Snapchat] *may share information about you as follows or as otherwise described in this Privacy Policy:* […] *With vendors, consultants and other service providers who need access to such information to carry out work on our behalf.*

We [Twitter] *engage service providers to perform functions and provide services to us in the United States and abroad. We may share your private personal information with such service providers subject to confidentiality obligations consistent with this Privacy Policy.*

We may share your Personally Identifiable Information with third party service providers to the extent that it is reasonably necessary to perform, improve or maintain the WhatsApp Service.

AFFILIATES

SNSs also share information with companies within the same group or which have common ownership:

We may share information we receive with businesses that are legally part of the same group of companies that Facebook is part of, or that become part of that group (often these companies are called affiliates).

We may share User Content and your information (including but not limited to, information from cookies, log files, device identifiers, location data, and usage data) with businesses that are legally part of the same group of companies that Instagram is part of, or that become part of that group ("Affiliates").

We may share your personal information with our LinkedIn affiliates (meaning entities controlled by, controlling or under common control with LinkedIn) outside of the LinkedIn entity that is your data controller (for example, LinkedIn Corporation may share your information with LinkedIn Ireland, or other LinkedIn operating entities) as necessary to provide the Services.

We [Ning] *may share some or all of your Personal Information with our parent company or any subsidiaries, joint venture partners, or other companies that we control or that are under common control with us (collectively, "affiliates"), in which case we will seek to require those affiliates to honor this Privacy Policy.*

PARTNER WEBSITES

Badoo links users' profiles with those on partner websites, although it is possible to opt out of this type of linking.

> *If you have registered on one of our partner websites, such as Facebook, your Profile will be available to all users of the Badoo social network whether via our site or our partners' websites.*

Instagram and hi5 have similar provisions:

> *We* [hi5] *may also share your clickstream information and information collected from cookies, pixel tags and local storage with our vendors and partners for the purposes of enhancing your user experience.*

> *A device identifier may deliver information to us* [Instagram] *or to a third party partner about how you browse and use the Service and may help us or others provide reports or personalized content and ads.*

The third party linked service is covered in the Myspace privacy policy, and use of the service authorises access to a very wide range of personal data:

> *When you choose to use a Third-Party Linked Service, you are authorizing the Myspace Services to share information generated by or available on the Myspace Services, including Personal Information, with Third-Party Linked Services, including, but not limited to: Full Name, Username, Profile URL, About Me, Profile Photo, Profile Cover, Unique Myspace Identifier (e.g., Myspace IDs), Stated Location, Gender, Age, Biographical or Demographic Information (e.g., professional title, college attended, etc.), Interests, Connections, Any Information that Is Publicly Available on the Myspace Services, Device location (including using GPS and longitude and latitude location information)*

USE OF CONTENT

Personal data or content from personal profiles may be used by the SNS:

> *We may use material that you post on the open access areas of Badoo in advertising and promotional materials on our partner sites and partner products.*

…when you post content to a LinkedIn Group that is open for public discussion, your content, including your name as the contributor, may be displayed in search engine results.

INFORMATION FROM THIRD PARTIES

Some policies also mentioned that personal data may be obtained from third parties:

Sometimes we get data from our affiliates or our advertising partners, customers and other third parties that helps us (or them) deliver ads, understand online activity, and generally make Facebook better.

… when you visit a third-party site that embeds LinkedIn professional plugins (such as "Share on LinkedIn" for publishers) we receive information that those pages have loaded in your web browser.

Myspace may also receive or collect certain information about you from third-party websites, platforms, or services in the event you choose to take advantage of various features that may allow you to integrate information on or to the Myspace Services.

We may also obtain information, including Personal Information, from third party sources. This includes certain Personal Information that may be provided to us through the installation and use of Ning Applications on third party web sites.

If you connect your Twitter account to your account on another service in order to cross-post between Twitter and that service, the other service may send us your registration or profile information on that service and other information that you authorize.

Facebook and Ning mention social plug-ins and other systems for sharing personal data with third parties. For instance Facebook's Instant Personalization is used to allow partner services to tailor services based on your Facebook profile and activity:

Instant personalization (sometimes also referred to as "Start now") is a way for Facebook to help partners (such as Bing and Rotten Tomatoes) on and off Facebook to create a more personalized and social experience for logged in users than a social plugin can offer. […] *When you visit a site or app using*

instant personalization, we provide the site or app with your User ID and your friend list (as well as your age range, locale, and gender).

As part of the sign-up process [Social Sign-in]*, the third party service will transmit certain information used to complete the new Member's profile page, including name, email address and, where available, profile photo, birthdate, gender and location.* [Ning]

PERSISTENCE

Some privacy policies are clear about the length of time that personal data is kept for and the reasons for doing so. When someone wishes to leave a social network and have their profiles removed from public view, there is usually a commitment to remove the data within a set period:

If you want to stop using your account it will be initially deactivated. […] *After the expiration of 30 days, your account is permanently deleted from Badoo. We delete photographs from our servers within 14 days from the expiration of the deactivation period while other information (such as contact history) is deleted by us over a longer timescale.*

When you delete your account, it is permanently deleted from Facebook. It typically takes about one month to delete an account, but some information may remain in backup copies and logs for up to 90 days.

After 30 days, we [Twitter] *begin the process of deleting your account from our systems, which can take up to a week.*

If you close your account(s), your information will generally be removed from the Service [LinkedIn] *within 24 hours.*

Instagram is less specific and simply refers to 'a reasonable time':

Following termination or deactivation of your account, Instagram, its Affiliates, or its Service Providers may retain information (including your profile information) and User Content for a commercially reasonable time for backup, archival, and/or audit purposes …

Two services, hi5 and Tagged (part of the same commercial group), declare that they reserve the right to keep data indefinitely.

The privacy policies also provide caveats about the operational requirements for keeping back-ups and the fact that they have no control over data which has been transferred or downloaded:

> *Warning: Even after you remove information from your profile or delete your account, copies of that information may still be viewable and/or accessed on the Internet to the extent such information has been previously shared with others, or copied or stored by other users or to the extent such information has been shared with search engines…* [Badoo]
>
> *Some of the things you do on Facebook aren't stored in your account, like posting to a group or sending someone a message (where your friend may still have a message you sent, even after you delete your account). That information remains after you delete your account.*
>
> *When you delete your IP Content or terminate your hi5 account, your IP Content is removed from the Site, you understand that we may maintain backup copies of the IP Content.*
>
> *Even after you cancel your Account or delete information or Profile Content, copies of some information from your Account or Profile(s) may remain viewable in circumstances where, for example, you have posted information on another User's Profile, shared information with a Third-Party Linked Service, or if another Visitor or Member copied, stored or shared your information or has a copy of the link to content you shared on the Myspace Services.*
>
> *Please note: Information you have shared with others (for example, through InMail, network updates, content sharing, or LinkedIn Groups) or that others have copied may also remain visible after you have closed your account or deleted the information from your own profile.*
>
> *…even if you update or remove Personal Information that you have provided to us* [Ning]*, your Personal Information may be retained in our backup files and archives for a reasonable period of time for legal purposes.*

Snapchat and WhatsApp operate on the principle of ephemeral data:

> *Once all recipients have viewed a Snap, we automatically delete the Snap from our servers and our Services are programmed to delete the Snap from the Snapchat app on the recipients' devices.*

> *If the recipient is not online, the undelivered message is held in WhatsApp's server until it can be delivered. If the message is undelivered for thirty (30) days, the undelivered message is deleted from our servers. Once a message has been delivered, it no longer resides on our servers.*

If someone dies there are provisions for either memorialising profiles or removing them altogether. Direct experience of this with LinkedIn and Facebook found that both SNSs required evidence of someone's death (for example, a published obituary or a death certificate) and the status of the person making the request (spouse, colleague, friend). They then attempt to contact the deceased person via their registered e-mail before removing the profile. This precaution helps to avoid malicious removal of profiles of living persons.

> *We may memorialize the account of a deceased person. When we memorialize an account, we keep the timeline on Facebook, but limit access and some features.*

Settings and User Options

Users have a degree of flexibility over what services or information they share and who can have access to it. This varies from profile and account settings, through to opt-outs (and sometimes opt-ins) to specific services. Table 12 summarises the settings and options that are mentioned in the privacy policies. A selection of these sites is investigated further in the technology section. The researcher signed on to a selection of services in order to gain access to the privacy settings of the SNSs.

TABLE 12 - USER-CONTROLLED PRIVACY SETTINGS

	Badoo	Facebook	Google	hi5	Instgram	LinkedIn	Myspace	Ning	Snapchat	Twitter	WhatsApp
Privacy settings											
Change privacy settings		✓			✓			✓			
Control who sees content	✓		✓	✓	✓		✓		✓	✓	
Search visibility		✓		✓							
Control collection of location info									✓		
Control access to sensitive info			✓								
Control communications from SNS	✓	✓		✓			✓	✓		✓	
Control third party use of personal data		✓				✓			✓		
Opt out of advertising			✓	✓		✓	✓	✓			

The first group of features in Table 12 relate to privacy settings. Only three privacy policies give instructions for changing general privacy settings, although most do give information about the options available for specific privacy settings. Controlling who sees content posted by a user or as part of their profile is a fundamental aspect of user management of privacy. Options include restricting to direct contacts, restricting to other members of the SNS or making the contents of a profile publicly discoverable. There was sometimes a warning that public profiles could not be restricted to any specific group. Facebook and hi5 both mentioned an option to prevent content posted by members becoming discoverable by search engines. Snapchat allows users to suppress location information attached to photos.

Many of the SNSs allow users to control the level and degree to which the provider communicates with them (usually via e-mail). This ranges from suppression of notifications of changes of status of contacts' sites, through to administrative and promotional e-mails from the SNS itself.

Third-party access to personal data has been the subject of much discussion (Shehab et al. 2012; Kontaxis et al. 2012; Solove 2011), however only three of the policies surveyed specifically mentioned this, possibly because of disagreements about what 'personal data' is.

Opting out of advertising by changing settings on the SNS site was mentioned in five of the policies. In one instance, the instruction was to use the IAB-run scheme to register individual preferences centrally. This was in addition to specific mentions of cookie suppression via browsers.

The individual policies varied by the degree to which they mentioned user options. For instance, WhatsApp makes no specific mention of user options, whereas Facebook provides instructions and detailed examples of how the user options work.

Safety Guidance and User Education

The Children's Online Privacy Protection Act 1998 (COPPA) in the United States is one of the driving forces for age restrictions on SNSs and this may be why 13 was specified as the minimum age of users in the privacy policies of: Instagram, Ning, Snapchat and Twitter. Myspace allows restricted use of their service between the ages of 13 and 18 with the additional protection of under-18s profiles not being browsable. WhatsApp has a minimum age requirement of 16 and Badoo and hi5 have a minimum age of 18. Badoo enforces this policy by retaining the e-mail addresses of those who have contravened its age policy to prevent repeated attempts to join the service. LinkedIn merely mentions that there is a minimum age requirement and refers users to their Terms and Conditions.

Many of the SNSs provide additional guidance to users about safety. Table 13 summarises the main aspects that were covered.

Table 13 - Safety guidelines

	Badoo	Facebook	Google	hi5	Instgram	LinkedIn	Myspace	Ning	Snapchat	Twitter	WhatsApp
Mention of additional safety guidance			✓				✓		✓		
Guidance on SNS's use of data			✓								
Guidance on technologies used		✓	✓								
Link safety guidance provided by an external organisation				✓				✓			
SNS safety site reviewed	✓	✓	✓			✓	✓		✓	✓	

Some of the safety policies have an educative role in that they alert users to the risks associated with misuse of personal data and steps that might be taken to ameliorate the risks. For instance Badoo makes the bald statement: "*When you post information about yourself or use the messaging, the amount of personal information you share is at your own risk."* It then goes on to identify some of the risks: *"These cookies do things like protect Badoo users from spam and fraud"*, *"…you may not use other users' information for commercial purposes, to spam, to harass, or to make unlawful threats"*, and *"We strongly urge you to periodically change your password to help reduce the risk of unauthorised access to your account information."*

The majority of SNS privacy policies identify fraud as a risk factor which they take active measures to protect against, but which also require the cooperation of users. The risk categories identified in the privacy policies shown in Table 14:

TABLE 14 - RISK CATEGORIES IDENTIFIED IN PRIVACY POLICIES

Risk	SNS privacy policy that mentions the risk
Fraud and unauthorised access	Badoo, Facebook, Google, Instagram, LinkedIn, Myspace, Twitter
Spam	Badoo
Safety, bodily harm or death	Facebook, Instagram, LinkedIn, Twitter
Harassment, bullying or infringement of human rights	Badoo, Google, LinkedIn, Ning
Illegal activity	Facebook, Instagram
Privacy violations	Facebook

COMPLIANCE

Although there has been recognition of the need for external validation of self-regulation and bodies such as the FTC monitor this regulation, there is still a problem with compliance. Well documented cases suggest that one problem is the *"discrepancy between the online service and their website privacy policy"* (Sherman 2012). Some of the SNS privacy policies acknowledge European compliance issues and make reference to them in various ways. Facebook and LinkedIn refer to their European or International subsidiaries for their non-US users:

> *Company Information: The website under www.facebook.com and the services on these pages are being offered to users outside of the U.S. and Canada by Facebook Ireland Ltd* […] *and is the data controller responsible for your personal information.*
>
> *If you live outside the U.S., LinkedIn Ireland controls your information.*

As residents of a European Union member country, UK users are afforded protection under the European Data Protection Directive (95/46/EC), enacted in the UK as the Data Protection Act 1998. The Directive requires minimum levels of protection for personal data of its citizens whether held in the EU or abroad. Potential trading restrictions can be applied to firms that do not comply with EU standards for data privacy. The U.S.-EU Safe Harbor framework, which was set up to protect EU citizens, is referred to in several privacy policies. The registrations are summarised in Table 15.

> *Facebook complies with the U.S.-EU and U.S.-Swiss Safe Harbor frameworks as set forth by the Department of Commerce regarding the collection, use, and retention of data from the European Union.*
>
> *LinkedIn complies with the U.S.-E.U. and U.S.-Swiss Safe Harbor Frameworks as set forth by the U.S. Department of Commerce regarding the collection, use, and retention of personal data from European Union member countries and Switzerland.*
>
> *Ning complies with the EU Safe Harbor framework as set forth by the Department of Commerce regarding the collection, use, and retention of data from the European Union. Ning has certified that it adheres to the Safe Harbor Privacy Principles of notice, choice, onward transfer, security, data integrity, access, and enforcement.*
>
> *Twitter complies with the U.S.-E.U. and U.S.-Swiss Safe Harbor Privacy Principles of notice, choice, onward transfer, security, data integrity, access, and enforcement.*

TABLE 15 - U.S.-EU SAFE HARBOR FRAMEWORK

Company	Status	Personal data processed
Facebook, Inc.	Current	Online, offline
Google Inc. and its wholly-owned U.S. subsidiaries	Current	Off-line, on-line, manually processed, human resources data
hi5 Recognition LLC	Current	Data from a client's HR system, limited in scope as agreed up on by the client
LinkedIn Corporation	Current	On-Line and off-line, HR data from the Company's EEA subsidiaries
MySpace, LLC	Not Current	On-line data
Ning, Inc.	Current	Online and offline
Twitter, Inc.	Current	Online and offline customer and employee data which may be manually processed

Facebook and LinkedIn also refer to TRUSTe and suggest this as a way of resolving privacy disputes, although Facebook is not a subscriber to the TRUSTe services (see Table 16):

> *As part of our* [Facebook's] *participation in the Safe Harbor program, we agree to resolve disputes you have with us in connection with our policies and practices through TRUSTe.*

> *We* [LinkedIn] *partner with TRUSTe because we take your privacy seriously and are committed to putting you and all of our Members first. TRUSTe certifies our compliance with the TRUSTe program and verifies our compliance with the US-EU and US-Swiss Safe Harbor programs. If you can't resolve a complaint through LinkedIn Customer Support, you may also contact TRUSTe.*

Ning refers to its compliance with the TRUSTe privacy seal:

> *Ning has been awarded TRUSTe's Privacy Seal signifying that our privacy policy and practices have been reviewed by TRUSTe for compliance with TRUSTe's program requirements and the TRUSTed Cloud Program Requirements including transparency, accountability and choice regarding the collection and use of your personal information.*

LinkedIn and Ning both specifically mentioned TRUSTe as an independent certification agent. The TRUSTe website provides the following information (August 2014):

TABLE 16 - SUBSCRIPTIONS TO TRUSTE SERVICES

SNS Provider	TRUSTe Service Used
LinkedIn	Dispute Resolution EU Safe Harbor Seal
Ning	EU Safe Harbor Seal Trusted Cloud

The EU Safe Harbor Seal operated by TRUSTe verifies compliance with the U.S.-EU Safe Harbor framework and replaces the self-certification provision in the framework. Dispute resolution is another service that directly relates to privacy protection and the provisions of the framework. LinkedIn refers users to this service in its privacy policy.

Google and Myspace commit, in general terms, to work with the appropriate regulatory authorities:

> *We* [Google] *regularly review our compliance with our Privacy Policy. We also adhere to several self regulatory frameworks. When we receive formal written complaints, we will contact the person who made the complaint to follow up. We work with the appropriate regulatory authorities, including local data protection authorities, to resolve any complaints regarding the transfer of personal data that we cannot resolve with our users directly.*
>
> *In certain countries, such as those of the European Union, you may have additional rights in respect of the processing of your Personal Information,*

including the existence of the right of access to and the right to rectify such information. [Myspace]

Some make the proviso that users may not have the same privacy protections as are prevalent in their own countries:

If you live in a country with data protection laws, the storage of your personal data may not provide you with the same protections as you enjoy in your country of residence. [Badoo]

Google processes personal information on our servers in many countries around the world. We may process your personal information on a server located outside the country where you live.

If you are located in the European Union or other regions with laws governing data collection and use that may differ from U.S. law, please note that we [Instagram] *may transfer information, including personal information, to a country and jurisdiction that does not have the same data protection laws as your jurisdiction.*

By providing information to Myspace, you do so in the knowledge that the laws applicable in your home country may not apply to protect your Personal Information or to your use of the Myspace Services.

The Ning Platform is hosted in the United States. If you use the Ning Platform from the European Union, or any other region with laws governing data collection and use that may differ from U.S. law, please note that you are transferring your personal data to the United States. The United States does not have the same data protection laws as the European Union and other regions.

Snapchat is based in the United States and the information we collect is governed by U.S. law.

Irrespective of which country you reside in or supply information from, you authorize Twitter to use your information in the United States and any other country where Twitter operates.

> *…you are transferring your personal information to the United States and you expressly consent to that transfer and consent to be governed by California law for these purposes* [WhatsApp]

Although some protections and rights are offered in a piecemeal fashion:

> *Users in certain jurisdictions are, in accordance with applicable law, entitled to exercise a right of access to personal information about themselves by asking for a copy of the information we* [Badoo] *hold about them (for which, where allowed by law, we may charge a small fee).*

DISCUSSION

DEGREES OF CLOSENESS

One key aspect of privacy that has arisen from this examination of privacy policies is control over who has access to personal information. There may be degrees of privacy, depending on the sensitivity of the information and how widely it is circulated. Some of the privacy policies attempt to define different types of audience for personal data or content posted by users.

A user may put information on his or her profile in the expectation that it is only available to their network, circle or friends. Other information may be intended for a more general audience, such as in LinkedIn where part of the purpose of joining is to see job and business opportunities. Some information may be made available to advertisers paying for access to personal data. There are different rings of access to personal data depending on its sensitivity, and the potential risk associated with providing access to it (Figure 11).

The further out from the centre, the less sensitive the information. For example, the inner circle might include financial information and personal data such as date of birth for verification purposes. The next level may include sensitive information such as personal interests and data required to complete transactions such as purchasing. This might be shared by close personal friends or direct service providers. A wider group is designated contacts, networks (LinkedIn), friends (Facebook) or circles (Google). A less personal group might be other members of group or discussions to which a user belongs. This identifies others with a similar common interest, but a user may not necessarily want members of one group to know which other groups he or she belongs to. For example, belonging to a church choir and a work-based football team may not have many synergies. The outer circle is other

members of the SNS service. Beyond the outer ring, is the general public, including non-members of the SNS.

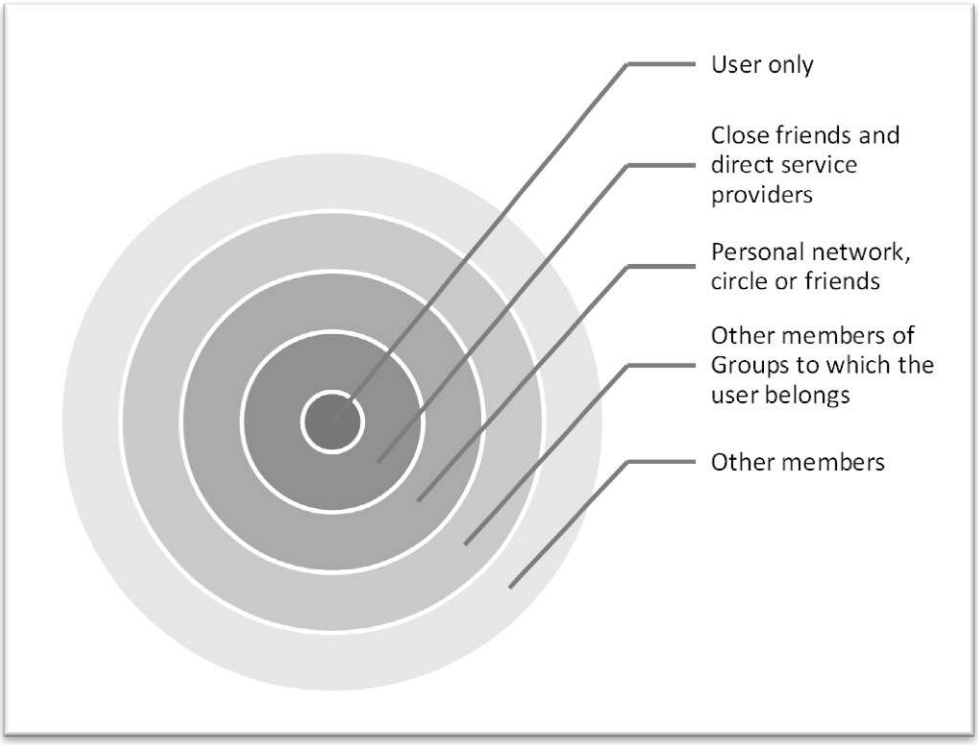

FIGURE 11 - DEGREES OF PRIVACY

Sometimes there is an assumption that personal information on profiles is only available to other members of the network. However for the general SNSs reviewed in this paper, this is not always the case. For instance, both Facebook and LinkedIn state that profiles are visible to external search engines and are therefore discoverable by non-members of the SNS.

REGULATION

Legislation represents one of the main modes of regulation of access to personal data, along with self-regulation as expressed in privacy policies. Some of the privacy policies refer to the legislation that might apply to the gathering, storage and processing of personal data. In the UK the main legislation regulating access to personal data is the Data Protection Act 1998. In an interview with an advisor from the Information Commissioner's Office, it was suggested that personal data on SNSs might be exempt under Sections 32 and 36 of the Act, discussed in Chapter 7.

One analysis of the application of the Act to SNSs suggests that the definition of a data controller is ambiguous when applied to SNSs. Comparison of legislation in Sweden, Germany,

Canada, Australia and the UK provide varying views on this. Some SNS providers explicitly state the scope of the definition of a data controller in relation to their own services (Garrie & Wong 2010).

Woods (2012) describes the tension between freedom of expression and privacy in SNSs. For instance the use of SNS postings by the media may conflict with the intention of users that content they post will not be widely distributed. Wider distribution of user-generated content could be seen as a tort of misuse of private information. It may also be seen as infringing on the right to a private life enshrined in the Human Rights Act 1998. The Press Complaints Commission (PCC 2012) excludes "*user-generated and non-edited material*" from the Code's remit in online publications.

SELF-REGULATION

Privacy policies are an expression of intent and provide one manifestation of self-regulation. However it is difficult to monitor the degree to which SNS providers adhere to their published policies. Several policies mentioned external agencies for monitoring and authenticating privacy policies by the Federal Trade Commission or by TRUSTe, the data privacy management company. To some extent policies are monitored by the public, in that if there is a major discrepancy between published policy and practice, this is likely to be publicised. There is certainly evidence that if users do not like changes to published policies this has generated considerable debate and protest (Helft & Wortham 2010; Kuzma 2011; Rubinstein & Good 2013). There is less evidence of protest associated with divergences from published policy and it is likely that that would have to be tested in court.

In the United States a review of the Federal Trade Commission (FTC) highlights unfair practices by Facebook, Google and Twitter (Hans 2012, p.167). In its settlement with Facebook, finalised in 2012 the report stated that:

> *The complaint alleged eight counts, including deceptive privacy settings, unfair and deceptive privacy changes, undisclosed dissemination of user information with third party advertisers, a deceptive "Verified Apps" program, and dissemination of user photos and videos.*

The review suggests that there has been a failure of self-regulation (Hans 2012, pp.197–198):

> *Years of self-regulation have failed to create an industry standard of privacy by design, opt-in sharing provisions, or other principles that would more*

> *effectively protect consumers. Web service providers consistently remain in the news for breaches involving user data. The status quo has been plainly insufficient in protecting user privacy, just as the current regulatory tools available to the FTC have been inadequate to ensure that companies are deterred from violating reasonable user privacy expectations.*

USER OPTIONS

The privacy policies do not document the full range of user options that are available and there is scope for development in this area. For instance, more attention could be spent on changing default settings to greater privacy. This study of policies goes hand-in-hand with direct investigation of live sites to document what options are available to users (investigated in Chapter 10).

As commentators have suggested user education, or 'norm change' may be the most effective way of protecting privacy in an online environment (Solove 2010, p.27):

> *The most effective solutions encourage norm change, and that occurs not just through the law but through increasing people's awareness of the consequences of their online speech.*

Privacy policies are a way of setting user expectations and in doing so they alert users to the possibilities for managing their own personal data in SNSs.

CHANGES SINCE THE LAST REVIEW

Although there have been some improvements since the last review of privacy policies (Haynes 2012), many policies are still opaque, repetitive and difficult to understand. One encouraging development is the greater emphasis on user education. The following services have separate online sites devoted to user safety:

Badoo
Facebook
Google
LinkedIn
Myspace
Snapchat
Twitter

Although it is difficult to be sure how effective privacy policies are on their own, they do have an important role in raising awareness and of setting the agenda for more general regulation.

CHAPTER 9 – SELF-REGULATION: DIGITAL ADVERTISING

WHY IS THERE A NEED FOR REGULATION?

Advertising forms the basis for the economic model used by social media providers. In return for a free service, customers provide personal information that can be used by online social networking services (SNSs) to profile their users. The SNSs sell access to targeted groups of customers to brands and advertisers. The more accurately they can profile a customer group, the greater the potential value to advertisers. Although direct ads are also placed on some social network pages, they are not targeted to the same degree as online behavioural advertising (OBA). Even though advertisers do not gather personally identifiable information, the act of categorising a user cookie can be seen as invasive. This touches on a wider problem of the blurring boundary between public and private information on SNSs and the tendency of services to move from 'Privacy by Default' to 'Disclosure by Default' (Strauß & Nentwich 2013).

The relationship between different agents in the advertising process is complex and continually evolving. Figure 12 represents a simplified schema of the different agents involved and provides a basis for understanding the descriptions of advertising transactions that follow (IABUK 2012).

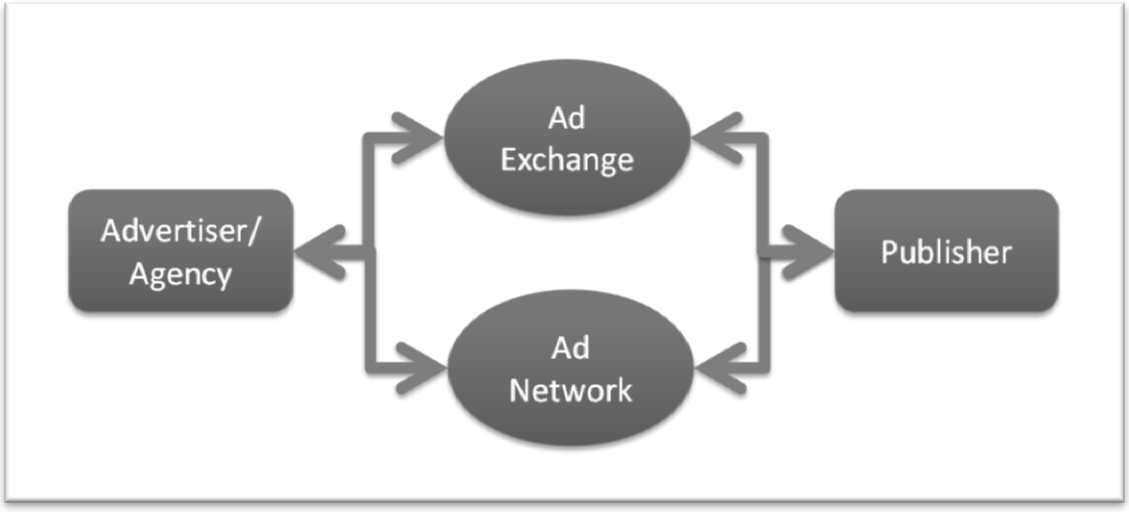

FIGURE 12 - AGENTS INVOLVED IN DELIVERING ONLINE ADS TO USERS

A publisher of a website may belong to an ad network with a number of other publishers. Through the ad network they serve targeted ads to advertisers (or their agents) in return for a fee. The ad exchange provides a mechanism for trading ads in bulk so that advertisers get the best price and publishers are able to sell available ad space on their websites.

This research sets out to consider the nature of risk associated with personal data on online social networking services (SNSs) and focuses specifically on the way in which this area is regulated. Consumers in effect make a contract with the SNS provider when they sign up to the service. In exchange for providing personal data about themselves, users are given access to a range of services and features via the social networking site. The service provider may sell on the personal data or access to personal behavioural data to third parties.

This study has highlighted the role of advertising in the dissemination and use of personal data. This was perceived as contributing to the risks faced by individual users of SNSs (Haynes 2014b). Regulation is one response to these perceived risks and this study identifies the regulatory modes or mechanisms that are used to mitigate that risk:

- Law
- Self-regulation
- Code
- Norms

WHAT IS OBA AND HOW DO SOCIAL NETWORKS FIT IN?

One of the reasons online advertising is effective, is the ability of brands to target their ads at specific audiences. The accuracy of a target group profile and membership can be improved by tracking the sites that a user accesses via their browser. This type of online behavioural advertising (OBA) is done by means of a cookie placed on the browser of a customer when they access a service (Figure 13) (Barth 2011). The cookie is detected on visits to other sites in the same ad network. The cookie is compared with an audience profile held on the ad network server, enabling the second site to deliver an appropriate ad to the user. The ad may also come via a trading system known as an 'ad exchange'. This operates like a stock exchange where the price of ads is determined by supply and demand. Each visit to a site will generate a cookie that captures details of that visit so that appropriate ads can be served to the user on future visits.

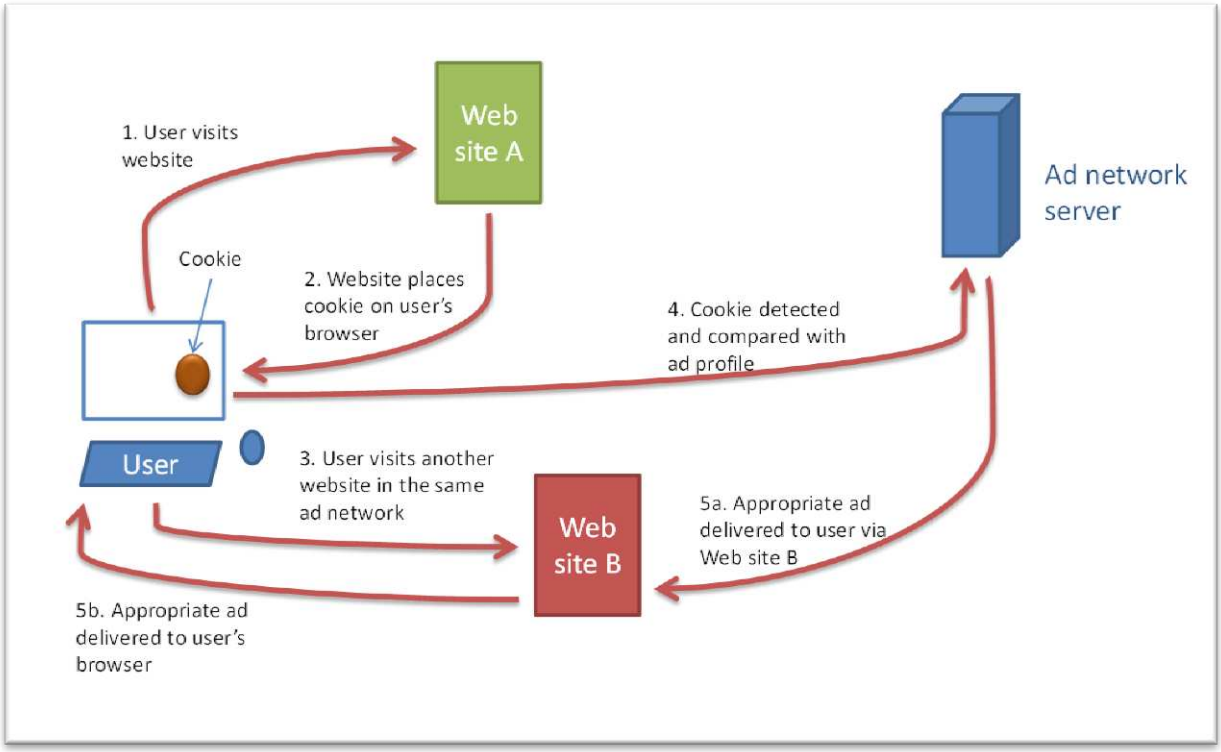

FIGURE 13 - HOW ONLINE BEHAVIOURAL ADVERTISING WORKS

The fact that this is not necessarily restricted to the social media site that placed the cookie is disturbing to many people. This may be one factor in the implementation of the ePrivacy Directive 2009.

Many social media also make APIs (application programming interfaces) available to interrogate personal profiles online. When consumers use social logins for instance, the service can interrogate the public profiles of the consumer and download relevant data to its own customer database. From the consumer's point of view social login might be seen as a convenient way of verifying identity. From the service provider's point of view it is a way of gathering sometimes quite sensitive personal data in order to target ads more accurately. Sensitive personal data includes religion, relationship status, details of postings and even personal details of the consumer's circle of contacts.

Once an advertiser has access to a consumer, either directly or via a social networking service, they can build up more detailed profiles by delivering short questionnaires, polls and games to find out more about the consumer and his or her attitudes (Gomez et al. 2009).

METHODOLOGY

A series of telephone and face-to-face interviews were held with representatives of the advertising industry and their regulators and were used to explore themes identified in an earlier survey and literature review (Haynes 2012). The purpose of the interviews was to build up a picture of the different modes of regulation used for online advertising via social media. They also provided an opportunity to consider ways of evaluating the relative effectiveness of regulation and possible future developments. Representatives of the following groups were interviewed:

- Self-regulatory bodies (online industry bodies, advertising bodies and professional bodies)
- Statutory regulators in the UK
- Academics working in the regulatory arena
- Technology companies providing technology-based solutions

Social media providers such as Facebook, Google, Twitter and LinkedIn were approached but were reluctant to take part in this research and are not included in this survey. Their privacy policies have been used instead as evidence of their views (Chapter 8). Most participants were happy for direct quotes to be attributed to them, and one has allowed for non-attributable quotes.

A series of semi-structured interviews (Appendix D) allowed for more open questioning and for the researcher to follow through information and ideas that arose during the session. The interviews were recorded and transcribed manually. A copy of the transcript was sent to each respondent for checking before a final version was imported into NVivo10 using for analysis using a combination of auto-coding and manual coding. The websites of the regulatory bodies were explored to verify information gathered during the interviews and to provide additional background information about advertising regulation.

INVESTIGATING ONLINE BEHAVIOURAL ADVERTISING

Online behavioural advertising (OBA) is the main revenue-generating activity of online SNSs. Several commentators have emphasised the benefits to the user of profiling and using cookies to ensure that ads are better targeted to users (Hammock & Rubin 2011). This was borne out during the interviews:

> *But there are huge benefits in being able to communicate and share things at the touch of a button, wherever, whenever, however. There are risks but I think the benefits massively outweigh the risks.*

Nick Stringer, IABUK

The advertising industry is keen to promote this message and this was emphasised in the public information website, www.youronlinechoices.eu produced by the European Interactive Digital Advertising Alliance (EDAA). For instance the 'About' page contains the following statement (EDAA 2015):

> *Online behavioural advertising (also known as interest-based advertising) is a way of serving advertisements on the websites you visit and making them more relevant to you and your interests.*

However one of the reasons for regulation is to address some of the actual, potential and perceived problems with OBA (Lynskey 2011). These include:

- Inaccurate profiling – leading to irrelevant ads, incorrect inferences about the consumer leading to disadvantageous terms or price discrimination for instance
- Profile disclosure – leaving the consumer open to ID theft and fraud
- Lack of transparency – difficult to hold profilers to account
- Unease caused by perception of surveillance or loss of privacy

The starting point for the investigation was to understand what regulation is already in place in the UK. The interviews and online investigation identified the following:

- Self-regulation of the advertising industry
- Privacy policies and EULAs (SNS self-regulation)
- Data protection legislation
- Technology and design
- User behaviour

SELF-REGULATION OF THE ADVERTISING INDUSTRY

The main response to online advertising up to now has been industry self-regulation. Much of the focus of self-regulation has been on codes of practice and published guidance. The Advertising Standards Authority features in most of the comments provided during interviews.

The social networks are slightly different and they self-regulate primarily via their privacy policies. Some take the view that SNSs are in fact providing advertising services:

> *But actually social networks work in a slightly different way in the sense that Facebook is the advertising network; Google is the advertising network; Twitter is the advertising network. So Twitter is serving you that advert, not a third party. The advertising industry's efforts have been more focused on the surrounding web browsing experience. They haven't gone into social networks, which is partly* [why] *the target case gets over-used. I do wonder if there's an element of – to what extent in the social networking world there is a worry that advertising is becoming too accurate.*
>
> Nick Pickles, Bigbrotherwatch

Nick Pickles of Bigbrotherwatch also suggested that:

> *The best regulators are still the consumers. Self-regulation beyond the law is in part driven by a fear of consumers.*

The European Advertising Standards Alliance issued a Best Practice Recommendation (BPR) on Online Behavioural Advertising in 2011. These recommendations are directed at self-regulatory organisations (SRO) such as the Advertising Standards Authority (ASA) in the UK. The BPR establishes three principles which have been incorporated into the UK's self-regulatory regime (Gray & Mills-Wade 2011):

1. Notice that OBA is being used
2. User Choice over OBA, including an opt-out mechanism
3. Sensitive Segmentation (i.e. children should not be targeted by OBA)

The advertising industry in the UK operates under a self-regulation regime that is funded by the industry through a levy. The Committee of Advertising Practice (CAP) produces a comprehensive code of practice which is monitored and enforced by the Advertising Standards Authority (ASA). The code of practice covers all aspects of advertising including online behavioural advertising. It is influenced by consumer protection legislation as well as the Data Protection Act 1998 and the ePrivacy Directive.

The ASA scheme is a complaints-driven system. When a member of the public registers a complaint, the ASA decides whether it is within scope and whether there is sufficient cause to

investigate. If an investigation takes place the ASA will adjudicate on whether to uphold the complaint or not. If a complaint is upheld, the advertiser has an opportunity to correct the problem before sanctions are applied. Sanctions include: adverse publicity; denial of licences to operate; and referral to other regulators or the Courts. The ASA depends to a large extent on 'naming and shaming' and publishes a database of adjudications as well as a list of non-compliant advertisers. They also apply pressure indirectly by for instance requesting search engine providers to remove links to websites of non-compliers.

The CAP Code has incorporated these principles into its rules on OBA, with Rules 31.1.1 and 31.1.2 specifically covering third parties. It is designed: *"To ensure that consumers are made aware of, and can exercise choice over, the collection and use of information for the purposes of OBA"*. The rules require: a clear notice on their website to consumers that data is being gathered for OBA along with a link to an opt-out mechanism; and a notice in or around the display ad. Rule 31.1.2 states (Committee of Advertising Practice 2014, p.122):

> *Third parties that use technology to collect and use information about all or substantially all websites that are visited by web users on a particular computer in order to deliver OBA to that computer must obtain explicit consent from web users before doing so.*

The ASA administers CAP rules on OBA, which *"apply to companies and organisations (referred to as "third parties") that engage in the collection and use of web viewing behaviour data for OBA via websites belonging to other companies or organisations"* (ASA 2013, p.3). During the first 6 months of operation of these rules ASA investigated 75 complaints. Many of the complaints were about failed attempts to opt out of OBA. Some of the complaints were based on misunderstandings of the provisions of the OBA rules (for example they do not cover right to privacy or human rights) or of the way in which OBA operates. None of the complainants were able to provide sufficient evidence to enable the ASA to carry forward their investigations. This represents a very low level of complaints, none of which were upheld in the first 6 months of operation of the new rules.

The ASA also monitored OBA 386 third party websites and noted 297 potential breaches of the CAP Code (rule 31.1.1) about prominent display of notices on *"collections and use of web viewing data"*. The 13 third parties falling within the UK's jurisdiction were informally notified of the potential breaches and asked to make appropriate changes to their websites. The ASA also examined a sample of ads from the 100 most popular sites visited by UK

consumers to see whether they complied with the CAP guidelines. Of 41 ads that were considered, six were thought to breach the rules (ASA 2013).

The Internet Advertisers Bureau UK (IABUK) provides self-regulation for the digital marketing industry and covers both media owners and technology companies. They maintain that self-regulation is better than statutory regulation because it can adapt more rapidly to changes in the market and technology, a view echoed in the United States (Christiansen 2011).

IABUK's main instrument of regulation is promotion of good practice by its members and the industry generally. It runs a kite mark scheme and has developed a public education website: www.yourchoiceonline.org as part of a European-wide initiative coordinated by the EDAA (European Digital Advertising Alliance) to educate the public about behavioural advertising. The website provides educational videos and background information about how behavioural advertising works as well as practical guidelines on how to protect privacy. There is also a link to an online utility which identifies the companies that provide online advertisements and allows users to change settings to opt out of a selection or all of the behavioural advertising provided by the companies. IABUK recognises the global nature of the advertising industry and therefore the need for an internationally coordinated approach to self-regulation. For instance, bodies such as W3C are currently working on a standard 'do not track' facility.

VIEWS ABOUT PRIVACY POLICIES

Another aspect of self-regulation is the policies and terms provided by the SNSs for their users. Having a policy that makes clear to users their rights and obligations is important. Ultimately they are enforced by contract law, although there may be some dispute about which jurisdiction holds sway.

The wider issue of informed consent raises the need for more transparent privacy policies and end user licence agreements (EULAs). There is also pressure on social media companies to adopt 'privacy by design' approaches to the development of their services. Some have also started responding to the need for greater transparency so that users can make informed decisions about how their personal data is used:

I think also the way in which social media systems work, they frequently change their terms and conditions, in the way they do things, in ways that can be quite confusing. They could do more to give people clearer choices and control mechanisms. I think people sometimes feel overwhelmed and again are lacking control over who really does see their information.

Iain Bourne, ICO

You have the fundamental starting point, data protection law is based on consent. I would argue that would also mean informed and meaningful consent and that you don't get informed and meaningful consent from asking someone to sign a privacy policy that is longer than the theory of relativity, that was written by lawyers for lawyers.

Nick Pickles, Bigbrotherwatch

VIEWS ABOUT DATA PROTECTION LEGISLATION

The UK's Data Protection Act 1998 is the principal statutory regulation governing personal data and most of the comments received were focused on this and on the principles that underpin the legislation. There are eight data protection principles which the Act encompasses including:

2. Personal data shall be obtained only for one or more specified and lawful purposes, and shall not be further processed in any manner incompatible with that purpose or those purposes.

3. Personal data shall be adequate, relevant and not excessive in relation to the purpose or purposes for which they are processed.

[…]

7. Appropriate technical and organisational measures shall be taken against unauthorised or unlawful processing of personal data and against accidental loss or destruction of, or damage to, personal data.

Principle 2 effectively means that personal data should not be 'repurposed'. If it is gathered for one purpose (providing access to social networking services) it should not be sold on or used for advertising, for instance.

Principle 3 means that the personal data should be adequate and relevant for the stated purpose – and that additional details of an individual should not be gathered if they are beyond the original stated purpose.

Principle 7 means that active measures should be taken to ensure the security and integrity of the personal data gathered by the SNS provider.

In May 2011 the ePrivacy Directive came into force in the UK specifically to address the issue of cookies. This requires all websites to make an explicit statement with an opt-out option if it places cookies on users' browsers. Since then a number of technology services have been developed to specifically address the desire of users to block cookies.

TECHNOLOGY AND DESIGN

Technology plays an increasingly important role in regulating access to personal data on social media. Cookie blocking software can be used as apps or add-ins on many browsers to identify and highlight cookies and allow users to block them selectively or comprehensively. The Ghostery software is one example of this. The EDAA also offers a utility via the www.youronlinechoices.eu website to block cookies.

Ad blocking software can also be installed to prevent pop-up ads from appearing on websites. This type of software is increasingly being adopted by users, a point that was noted:

> *Social media users and others are slowly but surely getting more assertive and more critical in terms of how people use their information and that they do find some forms of advertising intrusive. The fact that so many people are using ad blocking services is a bit of a worry I'd have thought if my business model relied on advertising.*
>
> Iain Bourne, ICO

Software designed for computer security and dealing with malware also includes anti-tracking technologies that delete or suppress cookies as one of a number of security features. There are also services such as tScheme and JanRain, which provide ways of sharing personal data in a more controlled manner, so that the user has a choice about who sees what.

USER BEHAVIOUR

Market responses and individual behaviour are both manifestations of individual attitudes about what is acceptable. In effect users regulate online advertising either by the decisions

they take as individuals (for instance to determine how much personal data they will share on social media), or by the cumulative effect of their collective behaviour (using alternative services if they do not like what's going on) .

The issue of user awareness has arisen in previous surveys and this is seen as one of the most effective ways of protecting users. As Guy Daines of CILIP (the Chartered Institute of Library and Information Professionals) says:

> *What we are really talking about here is information behaviour, in how people use information in their work and also in their personal life. It is about instilling the idea about responsible use.* […] *It's about changing the culture about being careful about privacy.*

It is also about personal responsibility:

> *The consumers who register on these sites don't read the small print about what they are allowing the brands to do with their data.*
>
> Russell Loarridge, JanRain

From the early days of SNSs commentators were calling for greater investment in user education, specifically in response to the use of advertising beacon technology (Gray et al. 2008). The collective behaviour of users can become regulation by the market, or 'mode' in Lessig's categorisation of regulatory modalities. For instance, if a significant proportion of users start to leave a service, the service may respond in a way that addresses user dissatisfaction. As one respondent put it:

> *the only reign in on those organisations is the damage to their brand if they did something stupid*
>
> Russell Loarridge, JanRain

Research by the Future of Privacy Forum in the United States *"to assess the communication efficacy of behavioral advertising disclosures on the web"* found that disclosure statements and icons increase the comfort of active internet users when confronted with ads on third-party websites (Hastak & Culnan 2010).

SHORTCOMINGS OF REGULATION

Respondents identified a number of areas where the existing regulatory instruments fell short of what is needed. This included:

- Lax legislation
- Poor enforcement
- Ignorance of legislation
- Controversy over the definition of 'personal data'
- Lack of awareness of options available to users

LAX LEGISLATION

The UK is seen in some quarters as being relatively lax in the way in which it has implemented the Data Protection Directive:

> *In our view the UK Data Protection Act is one of the weakest in Europe. Privacy International has done a comparator of all the implementations of this Act throughout Europe. We have even got a colour-coded map somewhere (it should be on the website). We did, a while back, an analysis of all the Data Protection Acts around Europe called 'Privacy and Human Rights' and there are reports on every European country published on our website.*
>
> Anna Fielder, Privacy international

POOR ENFORCEMENT

The ICO raised concerns about territoriality and the difficulty of enforcing the data protection legislation outside Europe. Iain Bourne believes that the attitude of search companies is changing, taking down harmful personal data or making it unsearchable and that this will accelerate when the General Data Protection Regulation 2012 comes into force. Since the interview, the Court of Justice of the European Union made a ruling which effectively brought forward this agenda (*Google Spain SL, Google Inc. v AEDP, Mario Costeja González*).

Others also identified problems with enforcement:

> *There is very little enforcement activity in this area through the system.*
>
> Malcolm Phillips, CAP

There have been very, very few prosecutions in the area of misuse of data or unauthorised sharing of data. Most prosecutions have been to do with telecommunications or miss-selling of PPI for instance, rather than misuse of personal data on social networks.

Peter Harris, BCS

There you would need to switch to the second approach, which would be enhanced enforcement of the rules. I know that this is something that has been discussed in the risk context at EU level. I don't see the relationship with risk myself, but you have regulators such as the ICO saying that because of their limited resources they need to take a more targeted approach to enforcement, which is entirely reasonable. They might even claim that is a more risk-based approach to enforcement – again, fine. But if you were only able to exercise your enforcement powers vis-à-vis a limited segment of data processing entities, then I think you have to have other methods to discipline other entities in the direct line of enforcement. There I think the backdrop of heavier fines would be quite useful.

Orla Lynskey, LSE

It is not effective at all. As I said earlier, I doubt it is legal, what's happening – as far as I know there are no test cases – because it is not informed consent. "When you log in we will take your geo position". Why? You are only supposed to collect data under the Data Protection Act if it's directly necessary and you've got permission to process it.

Anonymous Civil Servant

Even where enforcement does take place the argument is that the ICO is under-resourced to do this properly:

My problem is that the ICO is broken. I am not referring to the Office or the Commissioner, I am referring to their funding and their ability to investigate the law as it stands. […] *They almost invariably pick on public authorities, who aren't allowed to fight back, or small companies without the resources to fight back. You will rarely see the ICO take on a bank. They might get the FCA to take*

on a bank – using the sector specific regulators. I don't think legislation will achieve anything.

Toby Stevens, Information Privacy Group

Nick Pickles, at the time Director of Bigbrotherwatch, talked about differences in resources between large corporates and government:

If you look at the equality of arms between the individual member state data protection authorities and Facebook (a £500,000 fine from the ICO for Google is small change), whereas antitrust sanctions come with a much higher and more intimidating set of sanctions. I think that's partly driving it. There is a feeling from regulators that the existing sanctions regime is too weak.

He goes on to say:

Equally the data protection regulator is quite reticent to go after big corporates because they know that even if they are in the right they will lose three years in court and God knows how much they will have to lose in lawyers. There does seem to be a reluctance on the part of data protection authorities to enforce against the larger players. As a result they go after the public sector or very small players.

The limited territorial coverage of the DPA and the DPD is a major constraint for regulating services provided by global or international companies. The Information Commissioner's Office recognises that *"it can be difficult to link these multi-national, global companies to a particular territory"*. Peter Harris of the British Computer Society takes view that the effectiveness of legislation is likely to be limited because of territoriality issues and that alternative approaches are needed:

There are many countries where the intent to regulate is almost impossible. The only real regulation is almost by naming and shaming and a lot of publicity which discourages people from doing things which they don't appreciate the consequences of.

He goes on to cite the problem of service providers headquartered outside the European Union and the difficulty in applying regulations to them:

Regulation is incredibly difficult. One of the issues is for example, most of the social networks which are marketed within the European Union, which has strong data protection regulation, they are actually managed externally, from outside. I don't think EU law applies in Delaware [Headquarters for many major US corporations, including Apple and Google], *for example.*

Peter Harris, BCS

IGNORANCE OF LEGISLATION

The DPA is also limited by public ignorance. If people do not know their rights or know that there is a remedy for breaches of the data protection principles, they are unlikely to take action over breaches. Annual surveys of public awareness suggest that in 2013 40% of the UK population had heard of the Information Commissioner's Office. Only 33% of respondents in the ICO survey spontaneously mentioned the DPA when asked about laws concerning the handling of personal data, however when prompted this rose to 96% awareness of the DPA. This still means that most people would probably not know where to complain about a data breach, rendering the DPA ineffective for them. This perception is supported by other commentators:

For me, in terms of effectiveness, there are two things that could be improved. On the one hand, we could do more about user awareness. In recent months there has been an increased awareness even amongst the general population about the data protection rights. Those rights are, in my opinion, quite difficult to exercise for two reasons: it is difficult for individuals to identify the relevant actor when something goes wrong or when they want access to the data so that they can amend data; and then also you have a (sort of) de minimis problem.

Orla Lynskey, LSE

There is also ignorance about the application of the DPA:

> *The Data Protection Act places considerable responsibilities on data controllers, but even that term suggests that those people are capable of control. The issue is: who is the data controller, in terms of information that people may share in a chat room and they discover that it is public knowledge?*
>
> Martin Horrox, CIPR

CONTROVERSY OVER THE DEFINITION OF 'PERSONAL DATA'

Consideration of OBA throws up the question of what is defined as personal data. However the definition of personal data has been a topic of debate for some time (Millard & Hon 2012; Aldhouse 2013). In its response to the proposed DPR the Direct Marketing Association (2012) in the UK says:

> *The definition of personal data has been extended to include, in certain circumstances, online identifiers such as IP addresses. The EDPS believes that online identifiers should only be classified as personal data when there is a close relationship between the online identifier and an individual. This would mean, for example, that an IP address would only become personal data if the organisation holding the IP address knew some other piece of information to link it back to particular individual.*

However researchers have found ways to de-anonymize data used by advertisers on social networks by combining different data sets (Narayanan & Shmatikov 2009). In the view of some commentators this throws into question the whole 'pseudonymisation' approach (Information Commissioner's Office 2012). All this hinges on whether the data gathered in cookies is considered personal data.

There were also concerns about pseudonymised data and the proposed creation of a separate category in the legislation:

> *The introduction of pseudonymous data as a separate category of personal information – on the one hand it makes things like online advertising more certain, but when you apply that information to banking or medical data, it arguably weakens the protection systems. By creating this new category of*

information and acknowledging it has lesser protection you weaken protection around pseudonymous use of medical data.

In the UK we've been having this big debate about care.data. In our view that is personal data, whether it is pseudonymised or not, because if there is any risk of re-identification, then it remains personal information. So having that second category is a step backwards.

Nick Pickles, Bigbrotherwatch

LACK OF AWARENESS OF OPTIONS AVAILABLE TO USERS

Privacy International raises the concern that self-regulation is not very effective because users are not aware of the options or privacy settings available:

I don't think they are very effective really. I think they still target and profile to great detail and people still don't know about the fact that they are being targeted and profiled or how they can opt out of it. I don't know if there is any oversight of these self-regulatory codes, whether there are any complaints (where are the data?), how it is solved, is it self-policing. All these questions that are not answered or resolved. From our perspective effective data protection legislation is the best answer.

Anna Fielder, Privacy International

BENEFITS OF REGULATION

This review of the different models of regulation of advertising mediated by online SNSs throws into relief the relative merits of statutory regulation and self-regulation.

DATA PROTECTION ACT 1998 (DPA)

There was a degree of support for the data protection principles in the DPA by the advertising industry, because it sets a standard. The application of the eight data protection principles is seen as a flexible and appropriate way of protecting personal data:

Interestingly the 7 [sic] *key principles have proved relatively flexible and when combined with the Privacy and Electronic Communications Regulation (PECR) have created an environment in the UK where people are very comfortable.*

Chris Combemale, DMA

I think the principles are valid now. There is obviously a debate about 'the internet has moved on and we didn't have Facebook in 1998 so therefore we need to reform things', but the principles of the DPA exist well and good now.

Nick Stringer, IABUK

By setting a standard for good handling of personal data, the DPA is seen as being effective. It affects the behaviour of companies that use or collect personal data and helps them to make their privacy policies more transparent:

They achieve very little dealing with incidents. They probably have quite an effective deterrent effect for any legitimate business.

Toby Stevens, Information Privacy Group

I guess that because the kind of providers we are working with, who are tending to look at it [tScheme] *as a mainstream business (BT, the Post Office or the banks), they tend to have a healthy respect for the Data Protection Act and the Information Commissioner's Office and therefore go out of their way almost to do things in a way that they believe is correct.*

Richard Tevorah, tScheme

A good test of its effectiveness is to compare the situation with regions where there is no equivalent of Europe's data protection legislation:

However what I always compare this to is the counter-factual, which is the US. So a system where you have a total absence of rules governing the private sector and there you can see the effect the EU rules are actually having an impact in practice insofar as there is a minimum non-negotiable level of protection available for individuals, and companies will respect that.

Orla Lynskey, LSE

Risk reduction was not seen as an easily-measurable outcome that would allow investigators to assess the relative effectiveness of different regulatory modes such as the statutory legislation embodied in the DPA.

SELF-REGULATION

There was general support for self-regulation, not only from the advertising industry but also from other professional bodies:

> *…there is a lot of feeling that self-regulation and adverse publicity is stronger way of enforcing people's privacy than over-regulation which can clearly affect competitiveness.*
>
> Peter Harris, BCS

The Direct Marketing Association thought that self-regulation of the advertising industry works well in comparison with self-regulation of the press, for instance:

> *In marketing and advertising it works very well. The ASA has been particularly effective and when they announce their adjudication that people have over-stepped the mark, it tends to be quite well respected. There is an agreement that when the ASA has adjudicated against an ad, the media owners will not carry it, whether it's TV or press or whatever. Effectively it gets withdrawn voluntarily by the advertiser.*
>
> Chris Combemale, DMA

It was also seen as a preferable alternative to statutory regulation. For instance, the cryptographic industry has made the following case:

> *This is a new industry and it will kill the industry if you get too heavy-handed, and we don't quite know where the market is going. 'Why don't you work with us to come up with best practice processes and we, the industry, will self-regulate according to those standards. If you're happy those standards are appropriate, it should be good enough'. That was basically was the first tScheme as an industry body to police the self-regulation.*
>
> Richard Tevorah, tScheme

Toby Stevens suggests that self-regulation needs to be very focused, with clearly-defined regulatory mechanisms to be effective. He suggests that a bond paid by members of the industry would act as a way of rapidly responding to complaints and penalising offending companies:

Self-regulation is only going to work where it is sector specific. It only means anything if all the significant players in one space say 'Yeah, we'll sign up to this' […] *I've not seen any evidence of a self-regulation mechanism that has teeth. An effective self-regulation mechanism would be Identity assurance. For example, you have to deposit funds, a bond, like the travel industry, so that when there is a failure, the regulator can dip into that deposit to fix things. It's out of your hands. If you screw up, you lose the bond, and that allows instant recompense for affected parties, instant punishment for the culprits. It also creates a B2B mechanism (because this isn't just about consumers). Without that sort of mechanism I don't think self-regulation can be meaningful.*

The social network providers also regulate the way in which brands use personal data they have downloaded via APIs by applying contract law:

Each of the brands that we engage with has to have a contractual arrangement with those network service providers. It's basically the social networks protecting (in their language) their users' data. That's a euphemism for protecting the data that they've got and are already using for advertising. For example, Facebook will happily let you have access to the data on the customers that they have in the social network because they have published a bunch more [profiles] *than anyone else. You get 52 different elements of information on an individual from postings. But if you're going to use that information for advertising, they throw a flag on the play and they won't let you do it.* […] *There's no sense in taking the data from Facebook and putting the information into your database and using that for advertising, because you are contractually unable to do that.*

Russell Loarridge, JanRain

USER BEHAVIOUR

Others maintain that user behaviour is probably the best regulator:

To what extent in the social networking world is there a worry that advertising is becoming too accurate? Everyone always jokes when they get a Facebook advert that is trying to sell them a retirement product when they are 20; it's trying to sell them wigs or something bizarre – everyone always laughs: 'They

don't know everything about me; they're showing me all these wrong adverts'. It has got to a point where to what extent do they have to start deliberately showing me wrong adverts, because if they were right all the time, consumers would go 'Hang on a minute, they ***do*** *know everything about me'. It's the 'creepy line' (phrase used by Google and the NY Times). That creepy line is probably the better regulator than either self-regulation or legal regulation.*

(Nick Pickles, Bigbrotherwatch)

Iain Bourne of the ICO argued that regulation probably works because of the concern of SNS providers about loss of market through user behaviour:

That's getting very powerful. If you look at the way that groups of social media users grouped together to campaign against changes to privacy policies, changes to practices, it's really interesting. There is evidence that some of the social networking sites [are] *changing the way they do advertising, because of pressure from their own users, who are in a much better position through the services that these companies currently provide to find each other, to campaign, to join together.*

Iain Bourne, ICO

He went to on suggest that this was seen in the response of users to beacon advertising:

Do you remember the beacon advertising issue, where if you bought something on eBay, all your friends would be told what you had bought? That was very intrusive marketing which a lot of people really did not like. That was stopped largely through user pressure.

Future Responsibility for Regulation

A survey of LIS professionals suggested that primary responsibility for protection of personal data on social networks should lie with SNS providers (Haynes 2014b). When the interview respondents in this survey were asked who should be responsible for regulation, most felt that to some extent all stakeholders had a responsibility for what happens to personal data:

There are a lot of different elements here and we've all got our part to play. Regulators need to be a lot more connected to other consumer protection groups such as the OFT and people in advertising.

Iain Bourne, ICO

I think the responsibility has to be shared, in recognition of the different uses to which the data can be put, the different people who have access to the data and the responsibility of individuals who provide that data about themselves. I think that subject to consumer education it is possible to expect a certain level of individual responsibility among consumers for what data they make available. I think that, in as far as data is used for marketing purposes, then industry and industry regulators such as ours have a responsibility. And the platforms in as much as the platforms [such as Facebook] *employ that data directly for example to target advertising or to seek to sell products directly to the consumer.*

Malcolm Phillips, CAP

I think it is an obligation on everyone. There may be a role for legislation and statutory regulation. There may be a role for self-regulation. There is certainly a role for business and there is certainly a role for us as internet users. We all have a role to play, we all need to be accountable, we're all going to get benefits out of this, whether it is in advertising or in medical research. We all have a duty to have a responsibility.

Nick Stringer, IABUK

Government features in many responses, as do the tech, advertising and PR industries:

Tech companies. It is primarily the tech companies, the people who host this stuff have just got to find ways of being as responsive as they can. When there is evidence that stuff that is being posted is causing injury, distress, hurt – that's the bottom line.

Iain Bourne, ICO

There is a difference between regulation and policing and good practice. The regulation and policing – you do need a body like the Information Commissioner's Office in order to show that there is a penalty for not dealing with information responsibly.

Guy Daines, CILIP

There is an enormous role in government for educating and informing. And also for the regulators to inform. There is also a need for standards bodies to impose privacy standards on organisations certain software developments. But ultimately it's the commercial pressures on organisations to get the best marketing data, to make use of the information they have are very very strong.

Peter Harris, BCS

There was a strong feeling that individuals should have some responsibility for regulating access to their personal data, by their behaviour:

Clearly individuals have a responsibility to protect their own privacy if they understand what they are doing; to understand where it is being compromised.

Peter Harris, BCS

There has to be a greater acceptance of personal responsibility for dissemination of these data. It's only a tiny part of the problem, because I can give away my data quite legitimately and carefully and then the third party loses it or misuses it.

Toby Stevens, Information Privacy Group

Guy Daines of CILIP suggested that LIS professionals should have a role in educating users so that they can exercise greater control over what happens to their personal data:

At the end of the day the person who should be most aware of the risks they are taking should be the individual. They need to become an intelligent consumer, intelligent customer. That comes back to the role of library staff in actually enabling them or giving them some of the skills and evaluation framework to enable them to do that. If they are not interested in their own

personal reputation and safety, then I think it is going to be that much harder for other bodies, the Information Commissioner to do what they do.

CONCLUSION

REGULATORY EFFECTIVENESS

During the interviews there was general acknowledgement of the difficulties of trying to assess regulatory effectiveness. The ICO for instance has no formal mechanism for assessing regulatory effectiveness, although of course it does publish annual reports and commissions research on the impact of the DPA and other regulatory instruments within its remit. The DMA argues that if regulation is working, people won't notice and it is only when things go wrong that it comes to attention as in the case of regulation of the press:

> *You can tell when regulation is not working. When it's working people might take it for granted.* […] *When it doesn't go quite right as with the press and the phone hacking and where their self-regulatory structures were not enforcing because they were scared of Murdoch, they end up bringing themselves into disrepute.*
>
> Chris Combemale, DMA

Some suggested counting the number of reported incidents or complaints or prosecutions as measures of regulatory effectiveness:

> *The number of prosecutions, the number of complaints made against phone marketing are, I suppose some measures. It's not an area that's easy to measure.*
>
> Peter Harris, BCS

Or you could apply success measures used in other sectors:

> *We would assess regulatory effectiveness by the extent to which marketers agree to modify their practices in response to ASA adjudications, for example. Or the extent to which we are able to resolve satisfactorily the issues which complainants bring to us. We would regard those as the success measures for regulation. That would hold for this area and generally.*
>
> Malcolm Phillips, CAP

Country comparisons are another possible avenue, although, as was pointed out by one respondent, there are so many other factors that affect the situation in each country it is difficult to separate out the effect of regulation alone:

> *One way is to do country comparisons. Compare prosecutions, compare the role of the press, has implications on data protection and privacy. I think it is quite a challenge to try and measure it.*
>
> Peter Harris, BCS

> *I don't know whether you could do geographic comparisons. I don't know of methods beyond comparison that would help assess effectiveness.*
>
> Orla Lynskey, LSE

The BCS suggested surveys as a possible way forward for assessing regulatory effectiveness. CILIP referred to people's feelings of failure or injustice, which could be captured by a survey:

> *First of all you have got to get them aware of the problem, therefore you would expect the complaints and feelings of failure to be more. You would hope that what you were doing would bring that sense of failure, sense of injustice down, but it would be a difficult one to actually interpret*
>
> Guy Daines, CILIP

Cost benefit analysis is another possible avenue – although again it can be difficult to quantify this properly:

> *One of the questions I've asked is whether or not prior to the enactment of legislation we could have things like cost-back benefit analysis in the context of data protection.* […] *The question is again, this is similar to the problem you face in the risk-based approach, do you quantify the benefits? You could do this from a purely economic perspective, the benefits to business. But how do you quantify the benefits of data protection for the individual? I've seen very few people question that.* […] *it's more difficult to quantify the non-economic benefits.*
>
> Orla Lynskey, LSE

RISK REDUCTION

One of the purposes of regulation is to reduce risk to consumers. Nick Stringer of IABUK takes a wide view of the ways in which risks to users are reduced:

> *There are lots* [of ways of reducing risk to users]. *There's the EU initiative, also media literacy – informing people. There is also a very fast-moving technology-driven market that some people love and some people don't understand. We have a duty to inform people as to how it works that is meaningful.*

Mydex draws parallels with credit ratings. The ratings agencies are all members of a payments club and over the years have established a process for subject access requests and for fixing credit ratings. A similar process is evolving for data protection.

The idea of using risk as a means of assessing regulatory effectiveness had not been explored to any great depth by any of the respondents. Some rejected the idea outright or pointed out problems with quantifying and therefore measuring risk levels:

> *Some of the elements we are talking about here are just so difficult to quantify or to identify with certainty, that that's one of the downfalls of a risk-based approach when you are talking about something that is now a fundamental right.*
>
> Orla Lynskey, LSE

COMBINED APPROACH

The debate about regulating access to personal data has moved beyond privacy considerations. This demands a wider approach than relying on legislation alone. The interviews suggested that although legislation provided an environment that encourages responsible behaviour by advertisers, it is not enough in itself. It has the following shortcomings: lack of enforcement, territoriality, inflexibility, etc. As user behaviour evolves and new services are developed, legislation is in danger of lagging behind. Controversy about the 'right to be forgotten' is an example of legislation (and its enforcement) being out of step with current practice and market behaviour. Despite attempts to enforce the May 2014 decision by the European Court of Justice about the *Google Spain SL, Google Inc. v AEPD, Mario Costeja González* case, there is a growing realisation that this is unenforceable (Haynes 2014a; Powles & Singh 2014; Floridi 2014).

There is a strong argument for self-regulation, and not surprisingly, this is an approach favoured by the advertising industry. There is also a degree of support in other quarters, so that for instance the ICO sees this as a desirable adjunct to the legislation. However it is important that any self-regulatory system is credible and has independent oversight. The CAP Code of Practice and the enforcement regime administered by the ASA is considered an effective regulatory scheme that works largely through a combination of 'naming and shaming' and exclusion of non-compliant organisations from media and platforms provided by members of the industry. The ASA has avoided the failures of regulation represented by the finance sector (moral hazard), police complaints (poor compliance), and press complaints (regulatory capture). There is however a question about whether existing self-regulatory regimes effectively protect personal privacy. The self-regulatory regime in the advertising industry is primarily focused on the content of advertisements rather than the nuisance caused by advertisements and the potential breaches of privacy (ASA 2013).

Stauß and Nentwich (2013) see 'privacy by design' as a way of shifting SNSs toward privacy by default settings. There is a role for data protection agencies in scrutinising this but they also emphasise the need to raise user awareness, so that there is a multi-pronged approach to regulation. A study on attitudes to OBA in the Netherlands found that there was a great deal of ignorance about cookies and the authors concluded that there was a need to educate the public before legislation could be effectively applied (Smit et al. 2014).

These reflect Lessig's view that multiple modes of regulation working in concert are likely to be the most effective in the long term (Lessig 2006).

Further Research

Further work is needed in this area to compare the effect of advertising regulation on privacy and social networks outside Europe. Emerging economies such as the BRIC (Brazil, Russia, India, China) countries have large and growing populations of social network participants and which are targets for online advertising by global and regional brands.

This research has touched on ways in which regulation can be evaluated and there is an on-going need to develop some metrics that would allow comparison of different regulatory modes. There is a particular problem in quantifying the effect of regulation on privacy or in reducing risks to users, but the measures tend to be qualitative. Using measurements of risk might be one way of attaching a value to access to personal data and thereby providing a basis for comparing the most effective modes of regulation in this area.

Recruiting social network providers to this study proved difficult and this is an obvious area for further development. In particular it would be useful to understand their perspectives on current legislation and self-regulation and to explore what they consider to be appropriate future development of regulation in this area.

CHAPTER 10 – CODE AS A MEANS OF REGULATION

INTRODUCTION

OVERVIEW

'Code' is one of Lessig's (2006) four modes of regulation of the internet. In the context of SNSs, code is about the way in which a system is designed (or coded) and how that regulates access to personal data. Mostly this is about restricting access to specified groups. However there is a contradiction between this need on the part of users and the desire on the part of SNS providers to make personal profiles as visible as possible. Visibility helps in the recruitment of new members and it also means the provider can charge advertisers more for access to their customers.

This chapter examines the privacy settings of three major SNSs from the user's perspective: Facebook, LinkedIn, and Twitter. The case study protocol is described in Appendix E. These vignettes are used to build up a model of the types of personal data that is available on SNSs and who it is available to. It also provides an insight into the degree of control allowed to users.

The idea of privacy by design (PbD) is built into European legislation and is a tenet of the ICO's (2008) guidelines to data controllers. However there is a problem detecting whether PbD has been adopted. It is more of an approach than a measurable outcome. There is a danger that regulation can become too prescriptive and this approach of PbD offers an alternative way of embedding privacy in systems, rather than imposing processes on service providers.

Three technology providers, JanRain, MyDex and tScheme, were interviewed and their views feed into the analysis of options.

Finally, this chapter considers the effect of code as a regulatory approach and its effect on personal risk.

METHODOLOGIES USED BY OTHER RESEARCHERS

Facebook has been the focus of a lot of attention on privacy settings. As far back as 2005, the privacy policies of Facebook have been systematically surveyed. For instance, McKeon (2010) maps the increasing availability of personal data on Facebook default privacy settings between 2005 and 2010. A simple graphic interface demonstrates the dramatic change in the availability of personal information over this period. He defines five levels of availability of information. The information is available to:

You

Friends

Friends of Friends

All Facebook users

The entire Internet

He looks at the following classes of information: Wall posts, Photos, Likes, Name, Picture, Gender, Other Profile Data, Birthday, Friends, Contact Information, and Networks. Facebook's privacy policy has evolved in response to market pressure, changing legislative requirements, competition, and the perceived threat of new regulation. This process has continued as has the range of information types available.

In an investigation of the Facebook privacy settings of 400 randomly-selected information professionals in Turkey, Külcü & Henkoğlu (2014) developed a scoring systems for privacy protection based on the proportion of users that invoked a particular type of protection. This is an interesting methodology which could be applied to other countries.

This investigation of code considers the case of three contrasting SNS and explores the features from a user's perspective. This meant logging on to each service and making a record of the privacy options that are available to users. The presence or absence of a particular feature may in itself be significant in that it may reveal the provider's attitude to privacy and security of personal information. The settings available to users of the three services were compared in September 2014.

SETTINGS ON SNSS

Facebook, LinkedIn and Twitter form the basis of three vignettes to consider the way in which 'code'; is implemented as a regulatory approach. The attached CD-ROM (contents listed in Appendix F) contains screenshots that were captured during this research. These three services were chosen because they are widely used but have distinct purposes. Many people subscribe to all three. Facebook started primarily as a way for college students to socialise and this is reflected in its emphasis on social activities and on building up networks of friends. LinkedIn is a professional network directed towards career information and showcasing expertise. Twitter has much less personal data on its profiles but is a pervasive

communication medium used for personal networks, monitoring events in real time and as a leisure resource – also increasingly used as a campaigning and political resource.

PERSONAL PROFILES

All three vignettes allowed personal profiles to be edited. The choice that users make about how much information they put up on their personal profiles is probably the single most important decision that affects safety and security. If the data is not available on the personal profile, it may be much more difficult for someone (a criminal or a security agency) to get hold of. The three services vary considerably in the amount of personal data that is explicitly made visible via the personal profiles. However it is important to remember that a lot of personal data is also revealed by online activity such as postings, contacts made and favourites, which is dealt with later in this chapter.

Facebook requires users to use their real identity (as far as it is possible for Facebook to verify and enforce) even if they have a different public identity. This has prompted some debate and controversy about what someone's true identity is (BBC News 2014). Part of Facebook's success depends on links between users and online activity. The more that the service is able to encourage this, by disclosing changes to status and updates on the activities of individuals online, the more that a dialogue is created and the greater the opportunities for behavioural advertising to be delivered. In order to facilitate this, the default conditions tend to public disclosure. Each field of the profile can have its target audience changed independently. The main categories used are:

Me Only

Close Friends

Family

Friends

Friends of Friends

Public

Custom (e.g. by group membership or by location)

Another way of editing privacy settings is via the 'Edit My Profile' link on the user's Facebook Home page. Table 17 shows the default setting for different classes of data in the user

profile (September 2014). The majority of the profile data on Facebook defaults to 'Public'. Although making changes from the default is fairly straightforward, the user needs to know that this choice is available in the first place.

TABLE 17 - FACEBOOK DEFAULT AUDIENCE SETTINGS

Profile area	Default audience
Work and Education	Public
Places You've Lived	Public
Contact and Basic Info	Friends
Family and Relationships	Public
Details about You	Friends
Life Events	Generated automatically from other fields and so does not allow entry of new items

Facebook has also introduced a Privacy Check-up, which is accessible via the 'lock' icon at the top of the main screen, and shows the user their current privacy and security settings.

LinkedIn provides access to user Profile and Settings via an icon in the top menu. This invokes a password request for further security. A series of tabs allows settings including privacy settings to be adjusted from the default or previous settings:

- Profile
- Communications
- Groups, Companies & Applications
- Account

Under the Account tab the Activity Broadcast heading has a tick box to *"Let people know when you change your profile, make recommendations or follow companies"*. There is a warning underneath that this potentially alerts current employers if a user is seeking a new job. The Activity Feed defaults to 'Your connections' but can be changed to one of: Everyone, Your network (contacts of contacts), or Only you. What others see when they look at someone's profile reveals the name and headline, but can be changed to a generic heading such as 'Someone at City University London' or it could be totally anonymous. An individual's connections can also be made visible to only that individual. Normally they are visible to all his or her other connections by default.

The Communications tab allows users to select who can send invitations to connect, defaulting to anyone on LinkedIn. The privacy controls under 'Groups, Companies & Applications' allow users to turn on or off the data sharing with third party applications.

Managing advertising preferences comes under the Accounts tab. The default is to share non-personal data (i.e. aggregated or anonymised data) so that: *"LinkedIn may show me ads on third-party websites"*; and *"LinkedIn may show me ads based on third party data".*

Although there is limited personal data available on Twitter profiles – up to 160 characters – personal preferences and interests are revealed by tweets and locations from which tweets are made. It is possible to click on the profile of a tweeter and find out who is following them, who they are following, their tweets, photos and videos they have incorporated into tweets, as well as tweets they have 'favourited' and any lists that they subscribe to. Twitter is seen as a public domain and many people and organisations use it to raise their media profile. This is well and good for public figures, but the consequences of this level of visibility may not always be clear to many individual users. Like many social networks Twitter participates in behavioural advertising, and while individual identities are anonymised there can still be the perception of invasion of privacy.

Twitter allows users to change the default security and privacy settings via the Edit Profile page (accessed from the Settings icon in the top menu bar). The following settings can be adjusted:

Security

- Login verification
- Password reset

Privacy

- Photo tagging
- Tweet privacy
- Tweet location
- Discoverability
- Promoted content

It is also possible to request an archive of all the Tweets that an individual has made since first joining the service.

DISCOVERABILITY

Facebook posts are visible to the public as the default. These can be limited to specific audiences such as Friends, or Family at any time by clicking on the audience icon. From then on any subsequent posts will conform to the new setting until it is changed again.

No discoverability settings were detected on LinkedIn. This may be because part of the purpose of LinkedIn is to encourage business, employment and professional opportunities and limiting who can find a profile would be counter-productive.

Twitter allows users to opt out of having their tweets being searchable and visible by search engines that choose to comply, by using the 'protect my tweets' setting. This effectively makes the tweets only visible to followers. If this feature is invoked it then becomes possible to control who is a follower as new followers have to be approved.

Searches on Google routinely bring up Facebook and LinkedIn profiles and sometimes also Twitter profiles. The LinkedIn site provides a list of profiles fitting the name search and clicking on any one of them will reveal the full profile or as much of the profile as has been made public. If the searcher is recognised as a LinkedIn user, the level of connection is also indicated (e.g. 1st, 2nd and 3rd level connections). Facebook provides direct links to profiles with name, profile photo and favourites listed. Searchers are invited to log on for more services such as the opportunity to link up with the individual concerned. A list of alternative names is provided in case the selected profile is not the one sought.

LOCATION INFORMATION

Facebook users can change the location status of individual posts, including removal of the location altogether. The system can detect location on a mobile device and automatically add this to the post if required.

Although LinkedIn does track the IP addresses and locations of devices through which a user logs onto the system, it is not clear that this information is available to others. The user has the option of reviewing past location history from "Your active sessions - See where you're logged in" on the Settings page.

Twitter defaults to location not being revealed. Users have to opt in to revealing their location when they post a Tweet. It is also possible to retroactively remove the location tags of previously posted tweets. The level of specificity for location (e.g. 'Islington, London' or just 'London') can be adjusted as well.

ADS

All three services declared that they do not share personal information with third party organisations or sites. However third-party apps are able to harvest personal data or metadata attached to individual profiles. In the case of Facebook this is up to 52 fields of personal data including:

Posted By Username
Post
Tagged
Picture
Link Information (several fields)
Video
Type
Likes
Created Time
Updated Time
Comment Information (several fields)
Gender
Birthday
Location
Relationship Status
Bio
Religion
Hometown
Page Category
Page Likes

The three SNSs do share information with advertisers that is 'hashed' so that an individual is not immediately identifiable. This form of anonymising data allows the SNSs to gather information about user behaviour without breaching their commitment to not share personal data with third parties. This hashed data is used to deliver relevant ads to individuals based on their online behaviour.

Facebook allows users to supress ads that they do not like from their timelines by clicking the cross in the top right hand corner of the ad. The user is offered a choice of deleting the specific ad or all ads from that advertiser. Facebook then asks for feedback on the reason for choosing to suppress the ad. It is not clear whether this is taken into account in determining what ads are displayed in future.

LinkedIn allows users to opt out of ads from third party sites and ads based on third party data. This means that if a user's online behaviour is being tracked by a cookie placed on their browser by a third party, the user can opt out of that data being taken into account when ads are served on LinkedIn.

Twitter allows users to opt out of receiving tailored ads. This does not prevent ads being delivered to users, but it means that the ads are not based on tracking user behaviour and therefore not tailored to individual users.

AUDITING PRESENCE ON SNSS

Another Facebook feature, possibly introduced in response to the Europe vs Facebook campaign, is an archive function that allows a user to 'Download a copy of your Facebook data' (Shaw et al. 2012; europe-v-facebook.org 2014). This includes data that is not normally visible via a user's online profile, such as browsing activity and advertisements followed. The system takes a few seconds to create the archive and then sends the user an e-mail message with a link to access the archived data. It is password protected to prevent potentially sensitive personal data falling into the wrong hands. The data can be viewed as a text file or via a browser.

Twitter has a similar feature for downloading all the tweets that a user has made since joining the service.

No such feature was detected on LinkedIn.

EXTERNAL TECHNOLOGY SOLUTIONS

A number of research groups have been looking at ways of improving security. For example one group has worked on automating the definition of different social groups, which *"helps users set up their privacy policies automatically for their uploaded content"* (Squicciarini et al. 2014). Another group has looked at ways *"to develop a secure approach limited the access of users' data to the applications, and disclose only the attributes that the user consented"* (Shehab et al. 2012). One possible response is the launch of Simply Secure, a new service enabling ordinary users to access the technology in order to improve their security and privacy (Doctorow 2014).

AD BLOCKERS

A growing number of ad blockers are emerging on the market. It is a feature of several security software and anti-virus packages such as McAfee, Norton Security etc. Some search browsers such as Google Chrome also include ad blocking features that can be switched on. There are also 'Private Browsing' features such as that offered by Mozilla Firefox:

In a Private Browsing window, Firefox won't keep any browser history, search history, download history, web form history, cookies, or temporary internet files. However, files you download and bookmarks you make will be kept

There are also add-ins that can be integrated with commonly-used browsers (e.g. Mozilla Firefox, Google Chrome) such as:

Ad Killer
Adblock Edge
Adblock Plus
Ads No More
Blue Hell Firewall
Page Tweak
Smart Ads Blocker
Updated Ad Blocker

Ghostery is one example, which can be used to demonstrate the way in which ad blockers work. Ghostery was installed on Mozilla Firefox to monitor the trackers that are invoked when visiting the three test case SNSs. Table 18 shows the ad trackers detected by Ghostery on Twitter, Facebook and LinkedIn.

TABLE 18 - AD TRACKERS USED BY THREE SNSS

SNS service	Tracker detected with Ghostery
Twitter	DoubleClick, Google Analytics
Facebook	None
LinkedIn	DoubleClick, LinkedIn Ads, NetRatings SiteCensus, Quantcast, ScoreCard Research Beacon

There is also the EDAA's website and online facility for blocking trackers reported in Chapter 9. At the time of writing (September 2014) a browser add-in was under development to let browsers remember the last ad blocking settings, rather than having to reset them each time.

Effect of 'Code' on Personal Risk

Privacy by design and ad blockers are mostly focused on preventing personal data from being seen by potential abusers or misusers of data. The probability of data misuse occurring is reduced, but the impact of a data breach is unaffected by technology interventions. For instance, the design of a system to default to greater privacy when posting personal data means that there is likely to be less personal information available than if the default was for greater disclosure or disclosure to a wider group. Using ad blocking software to deny access to personal online behaviour also reduces the possibility of targeted ads being directed at a user. This could be an annoyance, because untargeted ads may still get through. However it does reduce the possibility of online behaviour being used for other purposes – such as surveillance by the security services and its attendant risks.

CHAPTER 11 – NORMS (CONSUMER MARKET AND USER RESPONSES)

INTRODUCTION

The purpose of this chapter is to consider ways in which users individually and collectively (the consumer market) regulate access to personal data. The chapter will look at evidence that user behaviour and market response have affected the way in which personal data is made available on SNSs, and how these are based on the norms that motivate behaviour.

The evidence used in this chapter is derived from the literature and an analysis of in-depth interviews with stakeholders with an insight into users and market behaviour. Analysis of privacy policies reveals the options available to users – backed up by online interactions and documented by the researcher (Appendix F). The results from a survey of LIS professionals in the UK (reported later in this chapter) also provides some insight into this mode of regulation.

In his description of 'norms' as a means of regulation Lessig (2006, p.124) uses some examples from the internet which now need to be updated in order to be applied to SNSs. He focuses on user behaviour and the ways in which groups can apply sanctions to individuals who transgress the boundaries (or norms) set by the group. While this is an important phenomenon it is also important to consider the way in which consumers apply sanctions to SNS operators as well.

> *Norms also regulate behavior in cyberspace. Talk about Democratic politics in the alt.knitting newsgroup, and you open yourself to flaming; "spoof" someone's identity in a MUD, and you may find yourself "toaded"; talk too much in a discussion list, and you are likely to be placed on a common bozo filter. In each case, a set of understandings constrain behavior, again through the threat of ex post sanctions imposed by a community.*

Westin (2003) discussed norms in terms of socially acceptable personal behaviour and ties this in with privacy. When an individual strays beyond accepted norms, a public response is required. It is no longer a private matter. In doing so he is describing the mechanism by which norms operate as a regulatory system.

Although some norms may be reflected in privacy policies and terms of use of SNSs, particularly in defining what is considered acceptable behaviour, different groups will have different standards of what they consider to be appropriate behaviour. Miller (2011) suggests that there are many different communities in Facebook, each with their own way of

working and norms of behaviour. Peter Harris of the British Computer Society illustrated the development of communities with their own languages and norms of behaviour:

> *…among teenagers, young people using social networks, completely re-defined what is meant by 'parents' 'friends' etc. They specified parents as close friends in order to bypass or twist the standard privacy settings on sites. So you will find that people are 'married' to people that they just happen to be friends with. Unless you understand the social background of people using these services you can completely misunderstand some of the profiles that are up there, but they* [young people] *understand what they are doing, but they are trying to subvert the intentions of what is up there.*

Norms are an emergent attribute of groups that are manifest in individual responses and observed in collective (or aggregated) behaviour. Norms may be manifest in responses to other users, for instance if someone strays beyond the boundaries of acceptable behaviour.

Taking up the Consumer View

Baldwin, Cave and Lodge (2012) talk about regulatory strategies that use 'market-harnessing' controls. They focus on regulating the markets though instruments based on legislation. They do not consider the effect that consumers have on regulating the behaviour of companies. The closest they come to regulation by the market (i.e. by consumers collectively) is the regulatory framework around competition. They go on to identify a number of laws in market regulation that effectively ensure consumer choice. The high cost of entry to the market can be applied to the case of SNS so that one cost model predominates (free at the point of use services and revenue generated through online behavioural advertising) making it very difficult for alternative cost models to develop. Large numbers of users and high market share are required to generate the level of income necessary to support and develop a new SNS. The other is the natural monopoly that early providers have gained for their markets. The way the market is structured, it becomes a matter of moving not only your own account but those of your friends in order to exercise choice in SNSs.

Users' concerns have been driven by a number of factors such as perceived risk, the desire for privacy, and concerns about surveillance by national security agencies. The Snowden revelations about the NSA's monitoring activity has highlighted the way in which users' social media activity is monitored and how this information is used for actions such as rendition,

arrest, and even targeting for assassination (Greenwald 2013). There are particular concerns about government agencies acting beyond the scope of the law, suspending due process.

The work of campaigning groups such as Privacy International, Bigbrotherwatch, and the Open Rights Group can be seen as both expressions of emerging social norms (around privacy and security) and as influencing public opinion about acceptable behaviour on the part of the authorities. However the impact is limited, as Anna Fielder of Privacy International pointed out in an interview:

> *We have a number of small digital rights organisations that really do very good work, but it is not clear how many ordinary people actually plug into them all.*

Pressure on SNS Providers

Wilson et al (2012, p.216) in their review of the literature on Facebook research suggested that users had a number of choices for managing risks:

> *(a) changing the level of privacy from the permeable default setting to a more private setting, such as friends-only status; (b) limiting the amount of personal information shared on Facebook; or (c) not acquiring a Facebook account*

Although these approaches represent significant choices by the market and therefore are potential means of modifying the behaviour of SNSs, this does not represent the whole range of possible responses. Existing users can also withdraw from Facebook (and other SNSs), or can request that material is taken down (with varying levels of success). Following the sustained campaign by an Austrian law student, Facebook now provides an automated service that allows users to download an archive of their transactions on Facebook (europe-v-facebook.org 2014).

User responses to privacy concerns have been considered by a number of researchers (Wills & Zeljkovic 2011; Willis 2014; Xu et al. 2013) and the range of responses to privacy concerns include the following:

- Cease using the SNS
- Change SNS
- Change privacy settings
- Change the amount of information disclosed
- Remove existing information from personal profiles

- Falsify personal information
- Adopt ad blocking technology and other privacy enhancing tools
- Campaigning

The last of these could be seen as an expression of social norms referred to by Lessig (2006). If a significant proportion of users begin to adopt these measures they may exert some influence on the SNS service and could be responsible for a 'norm' of acceptable behaviour concerning SNSs.

Nick Pickles (then Director of Bigbrotherwatch) acknowledged the effect of collective action and the sometimes contrary effect of legislation as an alternative mode of regulation:

> *There is a market pressure there, which sometimes if you regulate you take the market pressure in a way in which the consumers were not intending going.*

Attitudes of LIS Professionals

The content of this section has been published as an article in CILIP Update (Haynes 2014b).

A convenience survey of UK-based users of social networks was conducted. The 222 respondents were predominantly from the Library and Information Services (LIS) community (90% of respondents). This community was targeted because of their multiple roles as: users; intermediaries; and educators. The survey was conducted from mid-February to the end of April 2014.

The purpose of the survey was to find out what LIS professionals saw as being the main risks to adult users of online social networking services. It also explored attitudes to different regulatory approaches and considered who should have responsibility for implementing them.

The majority of respondents used Facebook, and Twitter on a daily basis and LinkedIn at least occasionally. Google+ was used by 31% of respondents. A total of 41 other social networking services were identified. Usage figures for the most commonly-used services can be seen in Table 19.

TABLE 19 - USAGE OF SNSS

Service	No. using the service at least weekly	No. using the service occasionally	Total No. of users	Proportion of total
Twitter	157	41	198	89%
Facebook	167	26	193	87%
LinkedIn	94	75	169	76%
Google+	23	46	69	31%
Pinterest	13	6	19	9%
Instagram	8	2	10	5%
Tumblr	7	2	9	4%

Total Respondents: 222

When it came to identifying risks faced by users, there was a clear ranking of risks in terms of their importance to individual respondents (Table 20).

TABLE 20 - RANKING OF RISKS

Item	Score	Overall Rank
Identity theft	1934	1
Strangers able to see sensitive personal details	1841	2
Targeting by advertisers	1575	3
Victim of fraud	1531	4
Discrimination by employer or potential employer	1443	5
Targeting by criminals (e.g. so that they can burgle your home while you are away)	1411	6
Friends, family or colleagues able to see sensitive personal details	1297	7
Cyber-bullying or harassment (including stalking)	1288	8
Targeting by official bodies or security agencies	980	9
Extortion or blackmail	628	10
Prosecution by authorities because of crime allegations	590	11
Physical violence or kidnapping	451	12

Total Respondents: 213

Users were most concerned about identity theft and about strangers being able to see sensitive personal details. Identity theft can itself expose users to other risks such as fraud and other forms of criminal targeting. Geo-location data was cited as another area of

potential exposure and risk. 'Strangers being able to see sensitive personal details' ranked much more highly than 'Friends, family and colleagues being able to see sensitive details'.

Some of the risks may have consequences that are more to do with social awkwardness or annoyance rather than loss of money or physical threat. For instance, targeting by advertisers may be irritating, particularly if users feel they are being spammed or cold-called as a result of existing personal data in their social media profiles.

Discrimination by employers or potential employers was ranked fifth. In addition several respondents identified discrimination by peers or by communities to which the user belongs as potential risks. This is often a result of data being used in a way that was not originally intended to make judgements that may be unfavourable to the user.

The cyber-bullying category includes: trolling; harassment; and "unwanted attention of a personal or sexual nature" – i.e. stalking. A new term, 'doxing' was used to indicate the threat of being 'outed' as a result of postings on social media, with potentially embarrassing outcomes or even reputational damage.

Loss of intellectual property was identified as a risk by several respondents who had experienced loss of images, designs, writing or other creative work as a result of postings on social media.

Several respondents expressed a concern about loss of control ranging from information overload to "sale of data you would not want sold". Some pointed out the problem of their personal data leaking through contacts with lower privacy settings than their own, or the consequences of being tagged by other users without their knowledge or explicit permission to do so.

When it came to who should have primary responsibility for protecting personal data on online social networks, 55% of respondents felt that it was the responsibility of the service providers, 26% thought it should be the users themselves, and 13% felt that it was the government's responsibility (Figure 14).

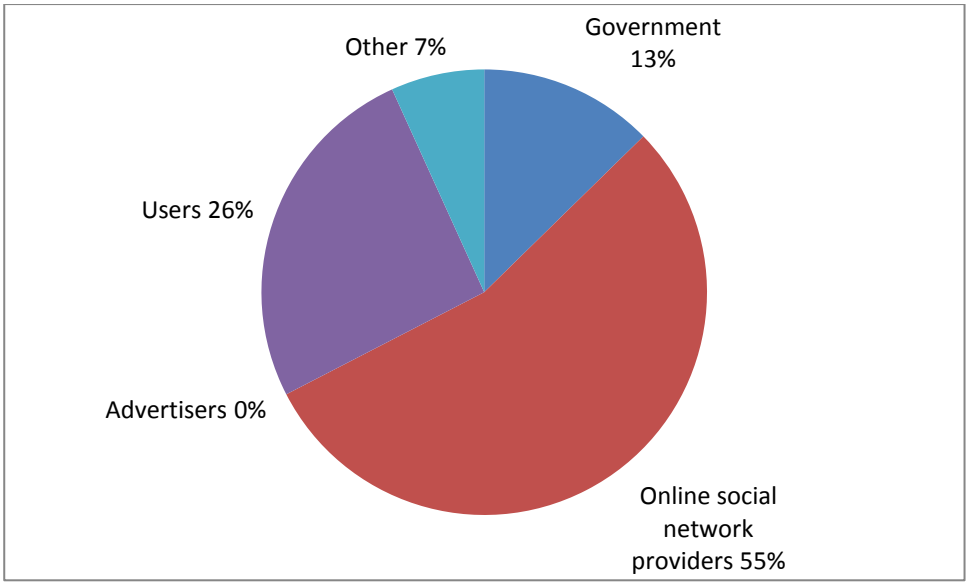

FIGURE 14 – VIEWS ON RESPONSIBILITY FOR PROTECTING PERSONAL DATA ON SNSS

There was no appreciable difference in attitudes between the different age groups or between males and females in the distribution of responses.

Several respondents felt that the users and the service providers should be jointly responsible for protecting personal data. One respondent suggested that a special body independent of 'big government' or the security agencies should be given responsibility for this area.

A more nuanced picture emerged from the degree to which respondents agreed or disagreed with a series of statements about current measures to regulate and protect access to personal data (Figure 15). There was agreement or strong agreement with the idea that personal profiles should only be visible to those people or groups that they specify (97% agree or strongly agree) and that SNSs should be set up with maximum privacy as the default setting (94% agree or strongly agree). Ninety-seven percent agreed or strongly agreed that SNSs should be designed with protection of personal data in mind.

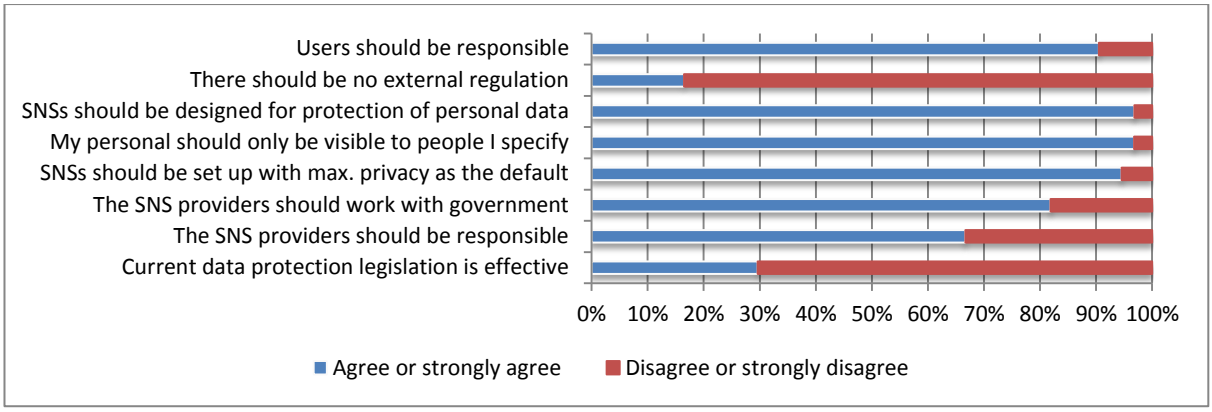

FIGURE 15 - ATTITUDES TO DIFFERENT REGULATORY MEASURES

Interestingly 91% agreed or strongly agreed that users should be responsible for their own online privacy, although in the previous question only 26% of respondents felt that users should be primarily responsible for protecting personal data on social networks.

Personal data should be regulated (84% of respondents), although 67% agreed or strongly agreed that SNS providers should be responsible for protecting personal data without government interference. A large majority (82% of respondents) felt that service providers should work with government to protect personal data. Other suggestions were that providers should be more transparent and make it clear which personal data is provided to third parties in what circumstances.

There is a clear picture emerging that the focus of effective protection of personal data should be the SNS providers themselves in the way in which the set up and operate their services and in the way in which they work with government. Only 30% of respondents thought that current legislation was an effective means of regulation.

User education and greater clarity in the guidelines provided to users on privacy settings were also seen as important measures. These included: more graphical interfaces to explain privacy settings, improved digital literacy on the part of users, and a partnership between government and providers to educate users.

Respondents wanted greater control over their data, including the requirement for providers to gain permission before gathering personal data or passing it on to advertisers or apps providers. They also wanted to be able to delete information or to archive the profile of a deceased former user. It was also felt that providers should not be able to change privacy settings and defaults without permission or at least prior notification.

The concern about the global nature of SNSs could be addressed by means of international agreements that govern the use of personal data. One view was that enforcement was the issue rather than the regulations that are already in place. One person suggested that there should be laws to prohibit government and security agencies from "spying on people".

Suggestions for controlling advertising ranged from outright prohibition to a system of more informed consent on the part of users.

The overall message from this survey was that current regulation is not particularly effective for protecting users of online social networking services against risk. Perhaps the greatest potential for mitigating risk is user education. This is one place where members of the LIS community have a particular role to play. This would help to ensure that users continue to benefit from the extensive range of features available on social networking services.

Norms Applied to SNS Providers

Following on from the survey, industry experts provided additional insight into norms that affect SNS services. For instance, Guy Daines of CILIP thought that individuals have an important role to play:

> *At the end of the day the person who should be most aware of the risks they are taking should be the individual. They need to become an intelligent consumer, intelligent customer.*

Users can respond collectively to 'unacceptable' behaviour by SNS providers either by organising pressure groups, or by changing the way they make use of those services. If a sufficiently large number of users change their online behaviour, this becomes a market effect. For example, in 2010 Facebook simplified its privacy policies in response to widespread consumer complaints about 'disclosure creep'. Successive changes to privacy settings had made initially private information more widely visible and this resulted in user information being disseminated in a way that was not envisaged when they originally disclosed it (Helft & Wortham 2010).

Chris Combemale of the DMA cites another example:

> *When Instagram wanted to use images that their customers had loaded and wanted to sell those images to advertisers and not share the royalties with the individual – of course it's just bonkers that someone would load their photos in*

> *a social media service, as a service to them and their friends and then you would sell those images to advertisers is completely incorrect – and it took about 48 hours of a major backlash of all their community rebelling against the idea. They had to back-track quite quickly. If Facebook goes too far, I think their members know them and they can move elsewhere now – is also effective in policing any company for customer service.*

Indeed according to Iain Bourne of the Information Commissioner's Office groups of users are now getting together to campaign for change:

> *If you look at the way that groups of social media users grouped together to campaign against changes to privacy policies, changes to practices, it's really interesting. There is evidence that some of the social networking sites* [are] *changing the way they do advertising, because of pressure from their own users, who are in a much better position through the services that these companies currently provide to find each other, to campaign, to join together.*

Norms of User Behaviour

Users can control what they reveal about themselves on online social media. An examination of the terms of service and privacy policies suggests that there is a minimum level of information required in order to register with a service. However many services encourage sharing of more information (particularly activity-based information) with wider groups of people in order to increase the connectivity and reach of the service. Pressure to do so establishes norms of what level of disclosure is considered reasonable.

Analysis of privacy policies, backed up by test registrations demonstrate a minimum level of personal information is required to enrol on an SNS (Table 21). In some cases there are checks to discourage falsification of personal data, although in most cases this is very difficult to control.

Of the 11 SNSs surveyed, 8 were available via a web browser (the remaining three needed to be downloaded as Apps) and required full name or first name and surname. The only exception was Badoo which only requires a first name. Most required an e-mail address (Facebook requires either an email address or a mobile phone number). Many also require the gender of the user and their location. Several also require users to select a username and/or a password. See also Table 9 on page 129.

TABLE 21 - REGISTRATION INFORMATION REQUIRED BY SNSS

SNS	Name	Gender	Birth date	Location	Email/Mobile	Other
Badoo	First	Yes	Yes	Yes	Email	Purpose
Facebook	First and Second	Yes	Yes		Email or Mobile	
Google	First and Second	Yes	Yes	Yes	Email and Mobile	
hi5	First and Second	Yes	Yes	Yes	Email	
Instgram					Email	
LinkedIn	First and Second				Email	
Myspace	First and Second	Yes	Yes		Email	
Ning	First and Second				Email	
Snapchat					Email	
Twitter	First and Second				Email	
Whatsapp					Email	

Users can exert pressure on their peers, especially when they step outside accepted norms of a group. As pointed out earlier in this chapter, there is no one set of standards that apply across social media or even within one SNS. This can work in a negative way, in that an individual perceived to be on the 'outside' can be ostracised, or in its most extreme form, bullied. Within a group certain types of behaviour might be considered unacceptable and the individual manifesting that behaviour may be subject to social controls. For example, using a personal account to promote a business service or to send spam e-mail to contacts could lead to being removed from people's online social networks, or being expelled from the service altogether. Some services have a mechanism that allows users to report material that is inappropriate or offensive. If there are sufficient complaints that material may be removed or the user barred altogether. Although the remedies are applied by the SNS provider, the resolution to act comes from peer pressure.

Well-publicised instances of trolling on Twitter and other media have revealed a deep revulsion at the misogynistic, threatening and unpleasant messages directed at prominent women such as Mary Beard (Professor of Classics at Cambridge University) and the journalist and feminist Caroline Criado-Perez (BBC News 2013b). However user responses are often not a satisfactory solution and, where available, legal remedies may be sought (Agate & Ledward 2013).

CONCLUSION

Work by the main UK regulators, the Information Commissioner's Office and the Advertising Standards Authority have emphasised the importance of user education. Analyses of user responses to privacy concerns have also consistently pointed to the need for user education. As one government official put it:

> *I don't argue that regulation should be what governments do, regulation is also what markets do. Regulation is also what good practice people do, because it's in everybody's interests. As the trends that we highlighted develop, more people will be aware of the choices they should be being given and things will change.*

This is a message that has been picked up by the Information Commissioner's Office in the UK and which forms a major strand of its strategy for improving data protection. One mode of operation of the ICO is to educate users and thereby modify the behaviour of users. This is one element in the ICO strategy to protect personal data on social media. The campaigns are aimed at children. Although this study has focused on adults using SNSs, the education of today's children will offer protection to tomorrow's adults (it is to be hoped). Guy Daines of CILIP drew an interesting parallel with swimming:

> *One of the key roles of library staff is to start to get people questioning* […] *asking the questions* […] *so that they are able to evaluate and to come to some form of risk assessment themselves. Using an analogy from the children's world but which could apply to the adult world. You have a fear of water – you can put in the lifeguards, you can give them buoyancy aids, but what you've got to do is to learn to swim. That is precisely what needs to happen within the online world.*

Indeed Peter Harris of the British Computer Society suggests that because of the problems of applying national law to international entities such as social media providers, social norms become one way of exercising some kind of regulation:

> *Naming and shaming is probably the strongest thing that people can do. There have been very, very few prosecutions in the area of misuse of data or unauthorised sharing of data. Most prosecutions have been to do with telecommunications or miss-selling of PPI for instance, rather than misuse of personal data on social networks.*

Nick Pickles (then Director of Bigbrotherwatch) summed up the effect of norms on other types of regulation:

> *Instagram is probably a good example. The best regulators are still the consumers. Self-regulation beyond the law is in part driven by a fear of consumers.*

He goes on to describe the power of the market (the collective manifestation of social norms):

> *In the middle you have the self-regulation which is trying to strike the balance of how far can we* [go]. *It's quite strange the self-regulatory function is actually "how far can we enforce our own terms of service, before customers start getting annoyed?". That's the benefit of having large networks rather than lots of small ones, is that if Facebook loses 2% of its user base, that's a critical problem for the business, because it could very quickly escalate.*

Evidence from the literature (Chapter 2) and from discussions with industry experts (Chapter 9 and this chapter) suggests that SNS providers respond to user norms and expectations. Where those norms are challenged (because of unilateral changes to privacy settings) users rebel and it is through the threat of legislation as well as potential loss of market share that SNS providers respond.

One of the ways of enhancing the effect of social norms is to educate users so that their expectations of privacy are reinforced and so that they learn appropriate ways of responding to requests for disclosure. User education forms an important part of the strategies of regulators such as the ICO and industry self-regulatory bodies such as IABUK.

SECTION IV. CONCLUSION

CHAPTER 12 – DISCUSSION OF RESEARCH RESULTS

INTRODUCTION

The purpose of this chapter is to evaluate the evidence to support the hypothesis that different regulatory modes have differing effects on risk to individual users of SNSs. It starts with a review of statutory regulation and its shortcomings as a sole means of regulation. It then considers self-regulation and the mechanisms by which SNSs regulate access to personal data. This includes the efforts made by the advertising industry. The design of SNS incorporating principles such as 'privacy by design' is considered as a regulatory mode. The chapter then considers the behaviour of individual users and collective behaviour manifest in the market and considers the evidence that this is being taken up as an effective means of regulating access to personal data. The second part of the chapter looks at personal risk as a means of evaluating regulatory effectiveness and presents the findings in Table 22.

STATUTORY REGULATION

One of the areas that this research set out to investigate was whether legislation alone was the most effective way of protecting users against the risks associated with use of SNSs. A detailed analysis of the legislation in Chapter 7 identified the EU's Data Protection Directive (95/46/EC) (DPD) and the UK's Data Protection Act 1998 (DPA) as most directly relevant to this issue. The Human Rights Act 1998 and legislation governing the online advertising industry were also considered.

The internet and services such as online SNSs are international in nature. The provider or website does not need to be located in the UK to deliver services to British users. This applies more generally to the European Union and its citizens. Providers may be beyond the jurisdiction of the countries in which their services are available and this presents a challenge to national authorities. In the case of the EU and US-based providers, an imperfect self-regulatory regime, the U.S.-EU Safe Harbor framework, has been put in place.

Another failing of the legislation is the limited resources available for enforcement. This has resulted in a paucity of rulings against SNS providers in breach of the privacy regulations.

Users have the option of taking providers to court if there has been a breach of trust – such as revealing personal details to third parties. This is a costly and time-consuming process and is therefore a remedy unlikely to be available to most users. The most notable case has been that of an Austrian law student who has taken Facebook to court and who has initiated a campaign 'Europe versus Facebook' (2014). To date there has been no similar initiative in

the UK. The DPA contains exemptions for journalistic use and for domestic use of personal data. It could be argued that these prevent users from protecting their privacy. This is compounded by the lack of clarity of who is a data controller in an SNS and what constitutes personal data. For instance, there has been discussion about whether the user or the service provider is the controller in the case of personal profiles on SNSs. There has also been an argument about whether aggregated data counts as personal data. The service providers tend to say not, and the European regulators are beginning to take the view that they are.

The other problem is the inconsistent way in which the DPD has been implemented across Europe. This allows SNS who do wish to be seen as compliant with the Directive to pick and choose the most friendly regulatory framework. This is a charge that has been made of Facebook Europe which is located in Ireland, a territory considered to have a less rigorous regime than the UK which itself is considered to have a relatively permissive regime compared with that of France or Germany.

The EU's General Data Protection Regulation 2012 addresses the issue of lack of consistency because it will be directly applicable across the European Union. Even so it does not fully address the problem of extra-territoriality nor that of enforcement which will still be the responsibility of national authorities. Enforcement has been strengthened so that the potential maximum fine has been increased to 10% of the company's global income, for instance. It also makes provision for users to move their accounts to other providers more easily, potentially strengthening the power of users.

Self-Regulation

An examination of the privacy policies of SNSs revealed that these are still evolving. Some of these changes have been well documented. For instance, the changing privacy settings of Facebook showed a move to greater disclosure over several years (McKeon 2010). This meant that personal data that was originally set at a default value of 'reveal to friends' gradually became publicly discoverable data via search engines. The privacy policies also revealed what data was gathered and who it is disclosed to. The policies were quite explicit about the minimum amount of personal data required to register as a user and this is borne out in the case studies. This included contact details and personal addresses. Behavioural and device data is also captured during SNS sessions. The behavioural data is aggregated and shared with advertising networks so that they can target different groups of users. This is a major revenue stream for most SNS providers. There was some dispute about whether this constitutes personal data or not. Another problem area for user data was the persistence of

personal data and the difficulty of removing it from an SNS. The privacy policies are declarations of intent and self-compliance is not independently audited (the TRUSTe scheme is an exception).

The two self-regulatory regimes investigated for this study were the ASA code of practice adopted by the advertising industry in the UK and the TRUSTe scheme in the United States. Groups such as the European Digital Advertising Alliance (EDAA) and IAB have also developed guidelines on online behavioural advertising. The main enforcement mechanisms are: name and shame; and peer-pressure. For instance, if an advertiser does not comply with the DAA guidelines, they can be ostracised by other advertising and media organisations, making it difficult for them to operate in the European market.

TRUSTe operates in the United States and is associated with the U.S.-EU Safe Harbor framework. This is not actively policed although in recent years the Federal Trade Commission, which is responsible for this agreement, has indicated that it might start to apply sanctions to non-compliant companies (Solove & Hartzog 2014). A study by Connolly (2008) suggested that many companies registered on the Safe Harbor scheme did not even comply with their own privacy policies.

The analysis of privacy policies and the investigation of the Advertising Standards Authority and the U.S.-EU Safe Harbor framework suggests that self-regulation does not afford much protection to users. However it does establish some principles of user rights and privacy and this is a starting point.

The Design of Systems and their Defaults (Code)

Several online SNSs were investigated by structured sessions with screen capture to document the options that users have for managing their personal data (Appendix F). There was a degree of discretion allowed to users in terms of what data they revealed to whom. The concept of successive rings of closeness in Figure 11 on page 151 is useful here because it provides a framework for comparison of the different audiences that personal data may be revealed to. Some less sensitive personal data (such as name) may be released to a wider group, whereas sensitive personal data such as religious or political views has additional protections and may only be viewed by a trusted 'inner circle' of contacts.

Privacy by design was developed as an approach in Canada and has been adopted by the ICO in the UK. The concept is that when systems are designed, privacy options should be built into them. From the user's perspective setting the defaults to greater privacy rather than greater

disclosure (the tendency of SNS at the moment) would do a great deal to improve the protection of personal data. Another recommendation by the ICO is that users should opt into having their personal data shared rather than opting out. Opting out is like a version of inertia selling; do nothing and your personal data is shared.

Users also resort to external technology solutions to protect their privacy. Perhaps the best developed category is ad blocking software. This type of software can be downloaded as an app or installed on a browser. They detect cookies and prevent them from being saved on the browser. This disrupts the tracking of user behaviour.

The use of code is primarily a voluntary activity on the part of SNSs although this may be subject to some pressure from users and legislators. It is very difficult to police this as it would require detailed analysis of the interactions of a range of user types with the SNS. The rise of a market for ad blocking software is perhaps the strongest evidence that the user options available on SNS do not give adequate protection to personal data.

NORMS

A survey of library and information service (LIS) professionals in the UK (Appendix C) was used as the initial instrument to discover the range of views that might be represented by users, intermediaries and instructors. From this perspective the LIS profession is a convenient way of gaining access to a range of perspectives. The survey revealed that although users accepted some responsibility for their own privacy the majority thought that the SNSs were primarily responsible for privacy.

Interviews with representatives of the regulators, the advertising industry and privacy campaigning organisations revealed a consensus that user education was key to improved safety and privacy online. As users become more educated, their expectations are likely to change and this in turn will influence the SNS providers. One analogy for user education is that it is like teaching children the highway code and instilling some kind of traffic sense. Rather than stopping the traffic or imposing severe speed limits (walking pace) to limit the damage caused by collisions with pedestrians, everyone is taught traffic safety awareness. The activities to inform and educate users about online safety could have a similar effect for SNS users.

REVISITING THE RISK MODEL

The review of risks created a model of five categories of risk faced by users of online SNSs. The relationship between risks and consequences in the detailed model in Figure 6 on page 84

can be summarised where the risks are described in terms of the outcomes that are expressed in terms of consequences to the user (Figure 16).

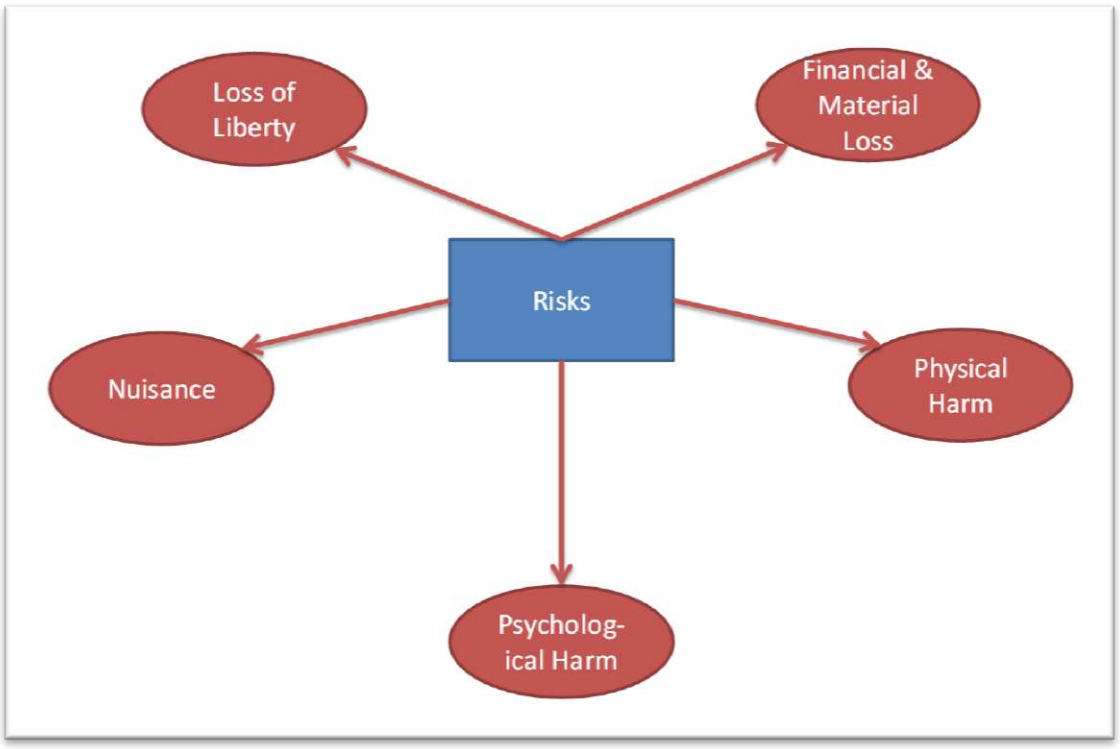

FIGURE 16 - RISKS AND CONSEQUENCES

NUISANCE

The investigation has highlighted advertising as one of the major areas of concern to users. These range from annoyance at being targeted with irrelevant ads, through to unease about use of personal information for targeting ads to concerns about the potential embarrassment if others become aware of what ads are directed at an individual.

PSYCHOLOGICAL HARM

Social media can be used to identity victims and to stalk them. The SNS can also be a vehicle for harassment, which in extreme cases can lead to self-harm. Cyber-bullying, stalking, trolling and harassment can all affect self-esteem and can undermine individual's confidence and ability to deal with everyday issues.

FINANCIAL AND MATERIAL LOSS

Identity theft from poorly protected social media profiles can make users vulnerable to financial loss through bank and credit card fraud, for instance. Disclosure of sensitive personal information can also affect job prospects (if there is evidence of unacceptable

behaviour) or insurance premiums (if there is evidence of ill-health). Revealing travel plans can make a user vulnerable to burglary.

LOSS OF LIBERTY

State authorities and security agencies sometimes act on the basis of what is on an individual's profile or by surveillance of their online behaviour. Examples might include boasting about criminal acts, visits to websites hosted by terrorist groups, or simply having a profile that corresponds to what the authorities are looking for. This might be quite a vague criterion that could be used to sweep up potential suspects.

PHYSICAL HARM

Two aspects of physical harm were discussed earlier: self-harm; and targeting for violence by others. There have been documented cases of individuals sheltering from domestic abuse being followed via their profiles and then being attacked in their new location (Thomas & Walport 2008). Criminals and gangs may also target individuals for kidnap or for retribution. Data revealed on the social profiles may identify them as potential targets and reveal sufficient information about location and habits to enable targeting for attack (Cox 2014).

EFFECT OF REGULATORY MODE ON PERSONAL RISK

A model of risk, which focused on the consequences to the user of a threatening event taking place on an SNS, has been presented. This produced a taxonomy of risk with five main categories. These risk categories have been used to evaluate the relative effectiveness of each of the four regulatory modes described in a model of regulation in Chapter 6. Table 22 on page 217 summarises the effects of each regulatory mode on individual risk.

LAW

In the UK the Data Protection Act 1998 (DPA) is the main statutory regulation in place for personal data on social networks. Early research suggested that SNSs might be exempt from the provisions of the DPA, on the basis that it is for domestic use. However that view has been questioned and it is now accepted by the ICO that social media do indeed fall within the remit of the DPA.

The EU's ePrivacy Directive, which came into effect in the UK in 2012, raised awareness of cookies and required all website providers to declare use of cookies and to provide users with an opt-out. Although this is theoretically backed by provision of enforcement, there is little indication that this is currently a priority for the ICO. Cookies are used by some advertisers to

track individual behaviour and thereby making deductions about the individual which can be used to target ads.

The DPA prohibits the reuse of personal data for purposes other than that for which it was gathered. It could be argued that data gathered to provide a SNS service should not be used for advertising, although many providers would then claim that this allowed for in the terms and conditions that users sign up to.

Extra-territoriality continues to be an issue. The fact that most of the SNS services reviewed are headquartered outside the EU means that they are beyond the control of the UK authorities. The result is that SNSs mostly ignore UK law.

The DPA does not specifically address the issue of individual misuse of personal data; its focus is on corporate behaviour. However other legislation such as incitement to racial hatred or sectarian hatred or human rights legislation can and has been used to prosecute individuals sending abusive messages to public figures via social media. This may have a deterrent effect on potential perpetrators.

Another abuse is appropriating and misusing data to make decisions that are detrimental to the subject. For instance, insurance premiums, refusal of credit, job selection can all be affected by unauthorised disclosure of personal information.

The law does not specifically protect against scams – and is always trying to play catch-up. It could also be argued that the DPA is ineffective in protecting against this risk, although other laws afford some protection. By extension the anti-discrimination laws protect against job discrimination and resulting financial loss.

The DPA has specific exemptions for security-related uses of personal data and for court-related information. In these instances the law does not protect users against the loss of liberty because of specific exemptions. In effect the law is used as an instrument to deny protection of personal data of individuals suspected of a crime and can even punish them before conviction, by depriving them of their liberty.

Although there are laws against physical violence, there is nothing specific in the DPA or other legislation to protect individual from the risk of physical harm. The *J19 v Facebook Ireland* case is an example where the plaintive wanted to publish details of a *"sectarian parade*

organiser" (the respondent) on a Republican Facebook website, which the court found would expose the respondent to the threat of physical harm.

Data protection law is probably neutral with respect to physical harm, because the provisions protecting individuals against harm are covered by other laws.

SELF-REGULATION

Annoyance of users is the area where self-regulation is most developed and evident. The advertising and public relations industries are governed by codes of practice with industry bodies in place to monitor the behaviour of advertisers and take appropriate action. Bodies such as the Advertising Standards Authority have extensive powers to get advertisers to change their behaviour. Self-regulation has a strong impact on the behaviour of advertisers and there is consistent focus on online behavioural advertising. There is pressure on advertising industry to behave responsibly as discussed in Chapter 9.

The SNS providers can intervene in extreme cases of bullying, but there is a marked reluctance for them to do so. This may be because they are wary of being accused of censoring the sites or of interfering with free speech. In this respect self-regulation is neutral.

SNS providers can block apps or users who abuse the system. Whether they are willing to take action is another matter. Commentary on recent events (such as the revelations about the NSA's PRISM programme) suggests that unless SNS and other social media providers rigorously apply their own standards of privacy in the face of pressure from governments, they can exacerbate the risk to users (Strauß & Nentwich 2013; Greenwald 2013; Bedi 2014; Morrison 2014; Witte 2013). On the other hand if SNS providers rigorously apply their own privacy policies they could afford some protection to users. The potential is for self-regulation to effectively mitigate this particular risk.

There is no evidence of self-regulation being an effective protection against physical harm. SNSs can remove threatening postings, but the practical difficulty of monitoring all postings militates against this. Some have a reporting function so that persistent offenders are identified and may eventually be excluded from the service. This may be because the number of reported incidents is low and the perception that this is not a priority for many users.

CODE

The design of SNSs and the privacy options they offer users affect the potential levels of risk to which users are exposed. Third-party technical solutions such as encryption and ad blocking software (to give two examples) can also act as preventative barriers to external threats. A variety of ad supressing software has entered the market recently and this is being adopted by significant numbers of users. The threat of widespread adoption of this software may have a moderating effect on some advertisers and service providers whose income is affected by the loss of potential target audiences. However there is a big security hole in systems such as Facebook because of the volume and type of personal information that is routinely made available to external apps or to third parties when social login is used. Without rigorous policing of third party apps and their providers, it is not possible to guarantee the security of any personal information transmitted to those organisations.

Some systems allow for unacceptable content to be flagged up for attention. This can work in two ways: either automatically removing or blocking offensive material; or by intervening manually when something has been flagged up. A variation on this approach is to have a minimum threshold of complaints that must be received before further action is taken, to defend against capricious and arbitrary decisions when someone is pursuing a vendetta. LinkedIn and other service providers have a mechanism for reporting unacceptable behaviour (such as spamming) which could include threatening behaviour.

The privacy options are important in limiting who has access to sensitive personal data. For instance, where credit card or other payment details are provided by a user, the card and transaction details must be kept securely and this is a result of the system architecture or 'code'. Other suggestions such as warning users when they are entering potentially sensitive personal data could be invoked and this is an aspect of code along with default opt-ins that require future development.

Some SNSs reportedly encrypt their messages to protect against interception. This approach is also available via a number of tools on the market that help users to encrypt messages or to restrict access to specific data. There is an emerging class of service where users store personal data on a server in such a way that allows them to control who or which services have access to the data.

The architecture of systems (whether it is the native SNS or an app associated with it) can restrict the amount of personal data that is available for mining or interception by intelligence

services, for instance. There is of course a potential arms race. As more sophisticated encryption becomes available, better techniques for breaking encryption codes emerge, and so on.

Some systems allow users to restrict sensitive personal information to defined groups of people. For example, the 'close friends and family' setting on Facebook can be used to limit access to very personal data to a smaller circle of contacts. The Google+ service works with the idea of user-defined circles, which can each be configured with appropriate viewing rights. This use of Code would tend to protect against stalking or direct physical threats.

NORMS

Norms are manifest in individual behaviour. Collectively that behaviour can be seen as a market response.

Norms can be seen at the most fundamental level of operation when people leave a service. If, for instance, they do not like the way in which their personal details are shared with third-party advertisers, they can vote with their feet.

There was some debate in a misogynistic trolling incident in the UK in 2013 about whether the subjects of attack should boycott services or whether they should out-face and eventually shame their attackers. Eventually there was such a groundswell of disgust that the trolls were put off (Agate & Ledward 2013). This was a good example of norms protecting against psychological harm.

User behaviour is a very important mechanism for protecting against financial and material loss. Teaching users to register with secure and unique passwords for instance will help to improve security and to protect against identity theft.

User behaviour will determine first of all what information is available, and to some extent how visible it is. Using privacy settings appropriately can improve personal security, although it is unlikely to be wholly effective against persistent and forensic investigation of individual profiles. Targeting of individuals by their online behaviour or by who they are connected to via SNSs can be avoided by modifying behaviour – but that restricts the utility of the service.

The use of non-language can mask the intent of groups of individuals. This has been reported in the literature about use of social media by teenagers, who in attempting to preserve their privacy against parental snooping have devised their own code or language to mislead external

users (Miller 2011). This could also be used by criminal conspirators and members of subversive groups.

In North America there is evidence to suggest that SNS providers respond to user concerns when there is the threat of user rebellion (Streitfeld & Perlroth 2012; Story & Stone 2007; Stone 2009).

Perhaps the best protection against physical harm is user behaviour. If users are aware that everything they post to social media is potentially in the public domain, it may moderate their behaviour. Some of the public education initiatives emphasise the importance of protecting personal safety and privacy. Although many of these initiatives are targeted at young people, they apply to all users. This raises the issue of the need to bring all users up to a level of 'traffic safety awareness'.

Using Personal Risk to Assess Regulation

The second major question that this research set out to address was whether personal risk can be used as a way of evaluating different modalities of regulation. A model of regulation was described in Chapter 6 and has been assessed in terms of their effect on the risk categories defined in Chapter 5.

Comparing different modes of regulation is difficult to do in a purely objective way, just as it is difficult to identify and assess risk where the categories are not well-established and there is no large corpus of statistical data to provide evidence of past events and consequences.

A grid of risk categories and regulatory modalities was used to create a risk-regulation matrix (Table 22). Evidence from the literature and documented events (such as court cases and actions taken by regulators) was used to determine whether a particular mode of regulation: a) mitigated against the risks to an individual; b) had little or no practical effect; or c) was detrimental to users with respect to the hazard being discussed. Each cell of the table indicates the mechanism by which regulation affects the risk category, whether it be positive(+), neutral (0) or negative (-).

TABLE 22 - EFFECT OF REGULATION ON RISK

RISK CATEGORY	Law[i]	Self-regulation	Code	Norms
Annoyance	**0** Cookie regulation DPA not enforced	**+** ASA and other regulatory bodies Codes of practice and industry bodies Effective for UK advertisers	**+** Ad supressing software Opt out of ads	**+** Leaving a service Ad blocking software Education by advertising industry
Psychological harm	**0** Not addressed by DPA legislation	**0** Infrequent intervention by SNS providers	**0** System allows users to flag up inappropriate content	**+** Flagging up unacceptable content Ostracism
Financial and material loss	**0** No addressed by DPA legislation Anti-discrimination legislation	**0** SNS providers can block scammers, but do not seem to very often	**+** A lot of data is automatically transferred to apps freely Better password control	**+** User education and behaviour geared to online safety
Loss of liberty	**-** DPA exemption for security and court cases	**-** Ineffective. All too often SNS providers capitulate to demands of security agencies	**+** Encryption of messages	**+** Deception by users Use of anti-language User boycotts
Physical harm	**0** Not covered by DPA May be covered by criminal law	**0** No evidence that SNSs are devoting much attention to this	**+** User-defined groups with access to control access to very personal data	**+** Safe behaviour to avoid risk

Key
+ = mitigates personal risk
0 = Neutral, or no effect on personal risk
- = worsens personal risk

[i] Primarily focused on the Data Protection Act 1998

The column for Law in Table 22 shows that the DPA exemption for national security actually exposes users to greater risks that their personal data will be used to their disadvantage. Personal data can be used to target and pinpoint a potential terrorist or criminal conspiracy leading to loss of liberty.

Self-regulation, particularly by the digital advertising industry, has a positive effect by reducing the annoyance to individuals caused by persistent targeted advertising. However self-regulation can be negative where SNS providers choose to comply with extra-legal data requests by security agencies.

Self-regulation is seen as most effective in the category that includes regulation of ads, a significant cause of annoyance to users. There may be a commercial pressure here because of the ease with which users can install ad-blocking software, denying the SNS providers of revenue generated from purchase of access to targeted groups of users. In some instances, such as handing over personal records to state authorities and security agencies, self-regulation is detrimental to users' interests because of its lack of accountability. Independent monitoring of self-regulation would be a necessary first step to strengthening the effectiveness of this mode of regulation.

Code is potentially effective in mitigating against all the risk categories although is not exploited to its fullest extent, despite active encouragement from the ICO to incorporate 'privacy by design' principles into the development of new systems. External technology solutions such as ad supressing software and encryption of data and messages are effective ways of managing annoyance and possibly also problems associated with financial and material loss and loss of liberty, through interception of messages and other personal data. Having the facility to control who sees what information is an important safety measure to protect against financial loss and against physical harm.

Code in the form of ad-blocking software has a positive effect by enabling users to opt out of ads and the consequent annoyance. Encryption and security software can help to protect against mining of personal data for fraud and thereby financial loss. Another positive effect is use of technology to protect against scrutiny of online activity by security agencies. This reduces the threat of loss of liberty. Countering this is the potential financial and material loss from inappropriate sharing of data gathered by Apps from websites. The design of SNSs to allow additional protection to sensitive personal data can limit the exposure to harassment, stalking and criminal targeting. This is achieved by allowing users to control who can see

what information on their profiles. Improved design and defaulting to greater privacy would enhance this protection further, although the economic drivers suggest that this is unlikely to happen. Any site that makes its money from selling on personal data is not going to discourage exchange of information.

Where norms include user behaviour, this is an effective strategy for mitigating all categories of risk. For instance, leaving a service or ostracising services and individuals who behave in a threatening or unacceptable manner can achieve a degree of protection. Learning a 'highway code' for use of social media helps to promote safe behaviour that avoids risks such as inadvertently publishing your home address, current location and movements and transmission of sensitive personal information to apps providers and other external sites.

Norms have a positive effect on all types of risk. Teaching users about online safety offers the possibility of greater choice of online behaviour. The collective behaviour of users becomes a market force.

The differing effects of legislation on each of the personal risk types suggests that personal risk can be used as a way of distinguishing between regulatory modes. It is possible that it may go beyond this by providing a measure of the relative effectiveness of each regulatory mode. The measures can only be comparative at this stage because there is an insufficient corpus of data available to accurately quantify these values.

It is important to set this in a wider context. Personal risk is not the only consideration that needs to be made when evaluating different regulatory modes. The effects of regulation on society should be taken into account. An individual threat such as loss of liberty may be a societal good, by reducing the risk of terrorist attacks, for instance. It is probably also desirable to take into account the effects of regulation on the operation of the market. Over-regulation may stifle innovation and lack of regulation may result in harm to individuals, society and ultimately to the markets themselves.

CHAPTER 13 – CONCLUSION

INTRODUCTION

SCOPE

This research set out to test the hypothesis that **law-based regulation alone is not the most effective way of protecting users against the possible risks associated with use of SNSs**. It also set out to test a second hypothesis that **risk to individual users can be used to compare the effectiveness of different modes of regulation**. In investigating these two hypotheses, first stated in Chapter 1, the following research questions were considered:

- What is the nature of regulation of access to personal data on online social networking services (SNSs)?
- What are the risks to users of having personal data on SNSs?
- How have law-makers responded to the inception and growth of SNSs?
- What effects do different regulatory methods have on risk to individuals?
- Is risk to users an effective method of comparing different modes of regulation?

In this concluding chapter each of the research questions is considered in turn before examining how the answers to these questions support the two hypotheses. The final section of the chapter identifies potential avenues of further research.

WHY REGULATE?

Before examining the different regulatory mechanisms at play in the social networking landscape it is important to understand what motivates regulation. Baldwin, Cave and Lodge (2012) in their work on the regulation of UK utilities have suggested that there are two main reasons for regulating: firstly to protect against market failure; and secondly to protect human rights and enhance social solidarity.

An example of market failure is the unequal bargaining power of consumers and service providers. This study found that consumers were not in a position to bargain about personal privacy because they did not have the expertise, nor do they have much power individually. This means that SNSs can adopt a 'take it or leave it' attitude. Another form of market failure is market monopoly. Although there are competitors, one service, Facebook, dominates the SNS market. Where there is a genuine choice, consumers (the market) are able to exercise some pressure on providers by moving their custom elsewhere (Rodrigues 2010). Wu (2010, p.304) in his study of the failure of regulation of the information industries in the United States focuses primarily on market regulation. He advocates *"maintaining a salutary distance*

between different functions in the information industry". He argues that the information industry is not a special case and that it is as prone to monopolies as other industries.

Chapter 4 of this thesis considers the other main reason for regulating SNSs, which is to protect human rights, including the right to privacy and the right to freedom from abuse and exploitation. The right to privacy is the basis for the UK's Data Protection Act 1998, discussed in Chapter 7. Lessig (2006, p.xv) asks the following questions in the preface of his book, 'Code is Law':

> *How do we protect liberty when the architectures of control are managed as much by the government as by the private sector? How do we assure privacy when the ether perpetually spies? How do we guarantee free thought when the push is to propertize every idea? How do we guarantee self-determination when the architectures of control are perpetually determined elsewhere?*

Blacklock (2003) in his study of regulation in the qualifications sector in England identified a number of problems with regulation. These built on Baldwin and Cave's (1999) work on regulation of utilities. Blacklock talks about regulatory capture where a regulator identifies too closely with the regulated industry. This may be because they come from that industry or because the possibility exists of being employed by that industry in future. For instance, where an individual moves from working for the regulator to employment as a legal counsel for an SNS provider, there is a real possibility that enforcement during their time as regulator might not have been as rigorous as it would otherwise have been.

Another problem is the limited territorial scope of regulation highlighted by Lessig (2006) and Wu (2010). The Internet is an international service and a service provider may be located outside the jurisdiction where its customers are based. For instance, Facebook Europe is incorporated in Ireland and is therefore not subject (so the argument goes) to British law – even if it deals with personal and confidential data relating to British citizens and residents.

BENEFITS AND DISADVANTAGES OF REGULATION

Earlier discussion in Chapter 6 suggested that managing risk was a benefit of regulation. This is despite the problems or defects that arise from the implementation of statutory regulations. Enormous effort and resources have been devoted to data protection regulation and this has facilitated a market where differences in privacy protection should not be a barrier to trade within the European Union. However this has called consumer protection into question, and especially the protection of individual privacy rights. There has been some concern about the

differences in the implementation of the Data Protection Directive (95/46/EC) in different European countries, an issue that the General Data Protection Regulation 2012 aims to address. Ireland, in its efforts to become a European hub for many US based firms has not only established a favourable tax regime, but is also considered to be lenient about data protection compared to countries like Germany (Zell 2014).

Regulation is expensive. It imposes costs on service providers and these costs are inevitably passed on to the consumer. This may be done indirectly by increasing the need for revenue generation through advertising for instance, which was discussed in Chapter 9. Or it could be by introduction of fees, although this seems unlikely given the currently established advertising-driven market. If badly designed, regulation can interfere with the efficient operation of the market and even the level of transparency. For instance, data protection legislation could prevent multi-national companies from operating efficiently if they were restricted by location in the way in which they could process customers' and employees' data. The legislation could create a trade barrier, an indirect tariff, on trade beyond the European Union's boundaries.

The disadvantages of regulation are compounded by the shortcomings in implementation of regulatory measures that do exist. One response is to consider whether regulation is necessary or desirable. Although there are significant benefits, the costs of achieving them may not make them worthwhile. It is difficult to enforce regulations on companies that are based outside the European Union. An analysis of privacy policies (Chapter 8) demonstrated that some providers do not consider themselves bound by European legislation and claim that by signing their Terms and Conditions users are agreeing to this. The other concern is that legislation could lull consumers into a false sense of security by believing that protections are in place, when in fact they are unenforceable.

Regulation can be justified in terms of the benefits to society, to government, to industry and to individuals. The People's Inquiry conducted by Demos emphasised the importance of transparency in the way in which individuals' data is used and regulated (Bradwell 2010):

> *The best defence against the inappropriate use of personal information and the harms associated with it is to make the use of personal information as democratic as possible. That requires a solid governance framework, including giving people the means to make meaningful, informed decisions about when and where to release information when it is in their power to do so.*

RESEARCH QUESTIONS ANSWERED

This consideration of the rule of regulation, its costs and benefits leads to the research questions that were posed at the start of this research.

RQ1 WHAT IS THE NATURE OF REGULATION OF ACCESS TO PERSONAL DATA ON ONLINE SOCIAL NETWORKING SERVICES (SNSS)?

Chapter 6 examined ways in which regulation applies to the online environment. Reidenberg (1998) in his model of regulation identified the following characteristics of internet regulation:

> *Lex Informatica has three sets of characteristics that are particularly valuable for establishing information policy and rule-making in an Information Society. First, technological rules do not rely on national borders. Second, Lex Informatica allows easy customization of rules with a variety of technical mechanisms. Finally, technological rules may also benefit from built-in self-enforcement and compliance-monitoring capabilities.*

This led to the development of a new model of regulation that, while acknowledging the importance of statutory legislation based on Law, considers other modes of regulation namely: Norms, Market and Code (Lessig 2006). The advantage of this approach is that it neatly addressed one of the major problems of legislation – its territorial limits. Lessig goes on to present strong arguments for the effect of Code on the operation of systems. He also considers the effect of collective norms on individual behaviour, particularly focusing on the power of ostracism to regulate online behaviour. This research extended the regulatory model to include the effect of collective user behaviour on service providers, as well as the effect of individual choices (Chapter 6).

Self-regulation is potentially a powerful instrument in the regulatory toolbox. The analysis of privacy policies in Chapter 8 indicates that there is an emerging consensus about protection of personal data and its use. There is some controversy about what is defined as personal data (Chapter 4), and this is an area that is likely to continue evolving.

The model of regulation developed for this study and described in Chapter 6 identified the following modes of regulation:

- Law – the statutory legislation at UK and European level that governs data protection and privacy

- Self-regulation – measures taken by SNS providers to protect personal data, and manifest in privacy policies and industry codes of practice
- Code – the way in which systems are designed and set up and technologies designed to protect privacy
- Norms – the values and behaviour of users, individually and collectively and their effect on SNS providers

RQ2 What are the Risks to Users of Having Personal Data on SNSs?

Baldwin, Cave and Lodge (2012, p.83) state that *"regulation can be seen as being inherently about the control of risks…".* This is a theme that has underpinned the approach taken by both the European Union and the UK Government in recent years (Better Regulation Commission 2006). Chapter 5 of this thesis identified the range of risks that an individual faces from using SNSs and the need to manage those risks.

After considering existing risk typologies, a set of risk categories based on consequences to users was developed. This helped to focus attention on the user outcomes rather than the causative events. This also provided a means of analysing and clarifying the complex relationships between cause and effect. The categories of risk identified were:

- Loss of liberty
- Nuisance
- Psychological harm
- Physical harm
- Financial and material loss

RQ3 How Have Law-Makers Responded to the Inception and Growth of SNSs?

Chapter 7 describes how law-makers have responded to the emergence of SNSs and the legislative framework in the UK for regulating access to personal data. This centres around the Data Protection Act 1998, based on the Data Protection Directive (95/46/EC). This legislation has its origins in the European Convention on Human Rights which identifies the right to respect for private and family life. One response to the global reach of SNSs has been the development of the U.S.-EU Safe Harbor framework. This self-regulatory scheme allows US-based companies (including SNSs) to process the personal data of EU citizens and residents in accordance with the Directive's principles. In 2014 the European Parliament passed the General Data Protection Regulation, which will apply directly to all EU member states.

Implementation and enforcement will be devolved to national authorities, such as the Information Commissioner's Office in the UK. The Regulation has a number of new provisions including 'the right to be forgotten' and the right to transfer personal SNS accounts to new providers.

RQ4 What Effects do Different Regulatory Methods have on Risk to Individuals?

Baldwin et al (2012, p.3) provide a threefold definition of regulation: *"As a specific set of commands"*; *"As deliberate state influence"*; and *"As all forms of social or economic influence"*. It encompasses statutory regulation via legislation, as well as other mechanisms that control or affect the behaviour of systems such as SNSs.

One of the purposes of regulation is to manage risk. Chapters 4 and 5 look at the nature of risk and who was affected when personal data is made available via SNSs. Although employers, service providers, and society in general are affected, the focus of this study has been on the individual subjects of the data being regulated. Individuals face the widest range of risk and are most affected by access to personal data.

The four modes of regulation defined in Chapter 6 were evaluated in terms of their effect on the five categories of risk defined in Chapter 5. Chapters 7-11 in Section III of this thesis consider each of the regulatory modes in turn (including two chapters on self-regulation). This is followed by a discussion in Chapter 12 which compares these modes of regulation and presents them in Table 22. It concluded that each mode of regulation ameliorates risk by decreasing the probability of occurrence or by decreasing the impact (consequence) of a risk event, none of them improves the outcome in all the risk categories. Some, such as 'Law' and 'Self-Regulation' could potentially expose users to increased risk of harm (such as imprisonment) by their policies of passing on personal data to security agencies.

EU legislation and the Data Protection Act 1998 set public expectations about the storage and use of personal data. Although the provisions of the law are quite extensive, the resources for enforcement are very limited. For most of the risk categories effectiveness depends on awareness of the individual rights and corporate responsibilities when it comes to handling personal data.

The review of self-regulation demonstrated was difficult to police, partly because of the lack of agreement on what constitutes personal data and partly because of the range of agents involved.

The way in which systems are designed (or Code) has a profound effect on usability of privacy options (for instance, whether they are there or not). The Information Commissioner's Office in the UK actively encourages system designers and service providers to adopt the 'privacy by design' approach. The main shortcoming is the imbalance between the power of the service providers and the regulatory authorities in the UK. It is entirely up to SNS providers to determine how they design their systems and trends towards greater disclosure seem to suggest that the main driver is the ability to generate revenue from advertisers. User experience is also a consideration and privacy may not be at the top of most users' priorities to begin with.

This brings us to the final category: Norms, as expressed by user behaviour. This is the one regulatory mode that can affect the outcome of most (but not all) the risk scenarios identified in Chapter 5. Individual consumers have the ultimate sanction of not using a service at all or of ceasing to use a service, if they do not like the conditions of use or the way in which the service providers conduct themselves. Users can also make choices (for the most part) about what information they divulge, who sees it and even whether it is real or made up. It therefore seems logical that more emphasis should be placed on user education (Christiansen 2011). This has been recognised increasingly by the service providers such as Facebook. Their privacy policies have become more accessible and interactive and the SNSs are beginning to provide tools that help users to easily monitor their privacy settings.

The digital advertising industry has also taken on user education. This is coordinated on a Europe-wide basis to provide consumers with consistent and helpful guides to safe usage of online systems (EDAA 2015). The authorities in the UK have developed or supported user education materials and guidelines to encourage and help people to use SNS more safely and to continue to benefit from the richness of features and possibilities that they provide (Information Commissioner's Office 2009; Get Safe Online Limited 2015).

RQ5 Is Risk to Users an Effective Method of Comparing Different Modes of Regulation?

Chapter 5 reviewed the risks and developed a typology based on the outcomes of risk events from the user's perspective. While this study aims to treat personal risk fully it was felt that it should not stray into areas such as child protection where special considerations occur, such as the meaning of consent, parental supervision and age restrictions on services. Other aspects such as intellectual property, censorship and national security are also potentially rich areas of future research.

In discussing the effect of different modes of regulation on personal risk, Chapter 12 demonstrated that each regulatory mode has different effects on the personal risk consequences identified in Chapter 5. This could be developed into a methodology that allows policy makers and legislators to compare and assess the relative effectiveness of different regulatory modes. This methodology could be used in a similar manner to privacy impact assessments.

TESTING THE HYPOTHESES

The research questions threw light on the two hypotheses that were proposed at the beginning of this thesis. They used empirical data collected and analysed during this investigation as well as evidence from the literature (Chapter 2).

H1 THE LAW IS NOT ENOUGH

The legislative approach to regulation has the benefit of being developed over several decades across Europe, with input from many different interest groups. In order to be effective, regulation usually has the force of law behind it. The law represents a codification of the public good, often contrary to the interests of the industries or enterprises that are being regulated. Legislation also provides a framework within which other forms of regulation can operate. This research demonstrated the relationship between self-regulation in the advertising industry and statutory regulation (Chapter 9).

The hypothesis that the law alone is not the most effective method of regulation begs the question: Is there a better way of regulating access to personal data on SNSs? The evidence gathered during this study highlighted the interdependence of different modes of regulation in this area, a factor acknowledged by the Information Commissioner's Office in its promotion of Privacy by Design (Code) and user education (Mode).

Chapter 7 identified a number of weaknesses in the legislative approach, namely:

- The problem of interpretation of exemptions
- The uneven implementation of the data protection directive across Europe
- The extra-territoriality and weakness of the self-regulatory U.S.-EU Safe Harbor framework
- The lack of means to rigorously enforce the existing statutory regulation

Although some of these points may be addressed by the General Data Protection Regulation 2012, the problem of variation in enforcement remains, as this will be the responsibility of the national authorities.

The Data Protection Act 1998 (and other national legislation) is limited in its territorial reach and therefore its applicability to SNS services based beyond the EU's borders. Alternative approaches such as persuading SNS providers to build more protection into their systems would help to overcome this problem.

The lack of resources available for enforcement is the other major limitation on the effectiveness of legislation. The ICO in the UK brings very few cases and tends to avoid cases with large corporates because they are likely to be difficult and expensive to pursue. Instead they rely on 'name and shame' tactics. Very few cases have been brought to the Courts by aggrieved individuals because of lack of awareness of their rights, the costs of bringing cases and the perceived likelihood of failure.

One of the benefits of legislation is that it sets standards and therefore the expectations of users. There is a view expressed by regulators, LIS professionals and experts in the field (Chapters 9 and 11) that users own at least some responsibility for safe use of social media and that some form of 'traffic sense' is required. Indeed Schmidt and Cohen (2013, p.54) maintain that *"security and privacy are a shared responsibility between companies, users and institutions"*. Few people would advocate banning vehicles from the roads because they are inherently dangerous. Instead there is a regime of statutory regulation that governs the behaviour of drivers and from a young age children are taught about road safety. It is possible to see parallels with the internet where some statutory regulation is required for the data controllers but a great deal of the responsibility for safety lies with the individual.

This suggests that while the law is very important for protecting personal data, it depends on other regulatory modes to be effective. Norms of users and the behaviour of providers are both affected by the law. Self-regulation also has a key role to play, a factor acknowledged by the Information Commissioner's Office in its efforts to promote PbD and encourage defaults to less disclosure. These factors all suggest that **law-based regulation alone is not the most effective way of protecting users against the risks associated with use of SNSs**.

H2 A Risk-Based Methodology for Evaluating Regulatory Effectiveness

One of the challenges of using risk as a way of comparing different modes of regulation is the difficulty first of all in defining risk, and then in measuring it. There are limitations to any

methodology and the principal issue with this approach is its qualitative nature. It is probably possible to say that one regulatory mode has a greater or lower impact on a specific category of personal risk than another. However it has not proven possible to quantify these relativities. One approach would be to assign a monetary value to each risk event or type of risk in the UK or even within the EU. For example, the monetary value of fines, damages, or compensation awarded following court cases could be added up. However there are too few cases for this approach to work. There is also an insufficient corpus of events to estimate probabilities of occurrence.

The second area of concern was the difficulty in separating out the effects of regulation from other environmental factors that affect personal risk. Cultural and national characteristics might have a more profound effect on personal risk than the regulatory mode. In some countries people might for instance be more concerned about freedom of speech and censorship, whereas in others privacy and security might be considered higher priorities.

A third objection to the hypothesis is the existence of other, better methods of comparing the effectiveness of different regulatory modes. Existing approaches focus on the effect of regulation on the market or on its impact on the industry being regulated (Baldwin et al. 2012; Serpell 2008). The socio-legal literature focuses on comparison of similar types of legislation in different countries (Radaelli 2010; Gray et al. 2008). Even though risk management has been one of the major justifications for regulation in Europe from the late 20th century onwards (Hutter 2005), the literature review in Chapter 2 revealed very little on the use of risk measurement as a method of evaluating regulation.

Chapter 6 identified four main modes of regulation of access to personal data on online social networking services (SNSs). Using personal risk as a perspective it was possible to distinguish different degrees of effectiveness of the different regulatory approaches. The risk perspective showed that there were shortcomings with each regulatory approach. For example, law does not always cover the territories in which the services are based, and there is an uneven implementation of legislation such as the Data Protection Directive across Europe. Self-regulation in the form of privacy policies is opaque to users and does not provide sufficient information for them to make informed decisions about their privacy settings. Code, the use of technology to protect personal data, is an emerging area but is very patchy and there is not a full suite of tools to monitor and evaluate SNSs. Finally, the users' behaviour (Mode) can be effective, if the infrastructure is in place for users to make meaningful decisions.

Risk has thrown light on these regulatory modes and provides some support for the second hypothesis that **risk to individual users can be used to compare the effectiveness of different modes of regulation**. However this approach needs further development if it is to be useful to legislators and policy-makers and the programme of future research described in the following section might help to develop the tools to do so.

Further Research

The model used to describe the different modes of regulation could provide a framework for a research programme to develop thinking on effective regulation of personal privacy in social media. The proposed programme would test this model of regulation further by extending it into other areas of activity, described below. It would also be the basis for comparative studies across different jurisdictions – possibly through collaborative research projects. The programme would have four main strands, corresponding to the four modes of regulation identified in this research:

Legislation – This study focused primarily on EU legislation as it applies in the UK. The regime is very different from the fragmented approach to privacy protection in the United States. Although there have been a number of comparative studies (Gray et al. 2008; Mendel et al. 2012; Schwartz 2013) there has been very little comparison between Europe, the US and emerging nations such as Brazil with large populations of SNS users. In the first instance a review of online privacy legislation in Brazil, Marco da Internet 2014, could form the basis of a comparison with the provisions of the EU's General Data Protection Directive 2012.

Self-regulation – The privacy policies of eleven major SNS providers available to UK users were investigated. Accessibility of these policies has been raised as an issue in a survey of information professionals (Haynes & Robinson 2015). A project could gather empirical data on the degree to which users consult privacy policies (tracking usage of the privacy pages of SNS providers, for instance) and what effect that has on their behaviour. There is also a real issue of the effectiveness of privacy policies in non English-speaking countries such as Brazil, which make extensive use of English-language SNSs. Attitudinal surveys of these groups could also point the way towards more effective user education and behaviour modification to avoid risk. Another aspect of self-regulation that warrants detailed investigation is the degree to which the privacy policies accurately reflect the choices that users face. Some aspects may be difficult to verify – such as whether SNS providers do actually destroy

personal data in the timeframes that they indicate in their policies – but others could be checked through a detailed auditing process. This would require the cooperation of the SNS providers.

Code – This research identified a number of tools that could be used to enhance privacy (Chapters 9 and 10). There is already considerable attention paid to Privacy by Design and this has been taken up as a principle by regulators in Europe and North America (Information Commissioner's Office 2008; Cavoukian 2012; Rubinstein & Good 2013). A comprehensive survey of external tools for protecting personal information privacy would complement this. For instance, data encryption tools, ad blockers and secure personal data centres are all areas that have seen recent growth in new products and services. A survey would identify emerging trends and might also help policy makers to identify gaps in the market and therefore inform research policy and investment.

Mode – The relationship between attitudes and behaviour warrants further work. The OXIS longitudinal surveys provide a useful starting point (Blank 2010). The proposed research would investigate the relationship between what people say (attitudes and reported behaviour) and what they do (actual behaviour). Attitudes to risk associated with personal data on social networks were considered in this research and the qualitative investigations identified a number of important risks and issues. This could be developed further by undertaking a national quantitative survey using a stratified sampling technique to see whether the attitudes of different demographic groups vary. Similar surveys could also be conducted in Brazil, North America, and other European Union countries, for comparison.

The other major strand of this research considered the role of risk to evaluate different modes of regulation. Social media are increasingly being used in the workplace in a number of ways:

- Use by the organisation for marketing and promotion
- Use by individuals in the workplace for private purposes
- Use by employees at home for private use

Future research could investigate the degree to which employers regulate usage of social media by employees, building on previous work (Wilkes 2011; Yokoyama & Sekiguchi 2014; Haynes 2011). The research would identify the risks that employers are trying to protect

themselves against such as: reputational damage; exposure to malware; security breaches; and lower staff productivity (Anon n.d.). It would also look at the extent to which these risks are addressed by internal information governance frameworks. The risk paradigm would be relevant in this context because one of the main purposes of information governance is to manage risk. Employers' social media policies have already been gathered and collated by Chris Boudreaux (2014) and this provides a rich resource for a preliminary investigation. This could be extended by surveying employers and their staff. A risk analysis could be tested as a way of evaluating information governance policies, in a similar way to the evaluation of regulatory modes in this research.

These proposed research strands would build on the two approaches tested in this thesis: the categorisation of regulation into four distinct modalities; and the use of risk as a way of evaluating regulatory effectiveness. The proposed research programme carries forward the conclusions of this thesis by providing an international perspective and by considering workplace issues. This fills a gap in existing research programmes such as those of the Berkman Centre (Harvard University 2015) and the Oxford Internet Institute (University of Oxford 2015).

CONTRIBUTION TO KNOWLEDGE

This research set out to explore the nature of regulation of access to personal data on SNSs by investigating the different modes of regulation and their effect on risk to individual users. One of the objectives of the research was to discover whether there was a way of comparing or evaluating the relative effectiveness of the different regulatory modes identified. The research was prompted by the rapid growth in the use of SNSs and the evolution and development of new services and facilities. Online social media provide considerable benefits to users which are free at the point of use. They facilitate greater interaction and create opportunities to connect with a wide range of people. The technology also presents significant opportunities for businesses to market to targeted audiences. However where these two interests collide, there is a perception of loss of control of personal data as well as a growing concern about misuse of personal data.

This research contributes to the on-going debate about personal security and privacy in the context of online social media and has made the following unique contributions to the field:

- It has developed an approach to assessment of regulatory effectiveness based on risk to individuals (Chapter 12). This method could be extended to cover wider social

media or to consider use of SNSs in the workplace. Although one of the acknowledged purposes of regulation is to manage risk, there was a gap in the reported research about using risk to evaluate regulatory effectiveness.

- A new typology of risk was developed (Figure 6) based on the consequences of risk events from the perspective of the individual user. This model was developed because existing models were either too general (Swedlow et al. 2009) or where they applied to SNSs were very partial in their coverage.
- The regulatory landscape was investigated using Lessig's (2006) model of four modes of regulation as the starting-point. Lessig's model predates the rise of social media and does not take full account of self-regulation. This has become a much more prominent feature of regulation and deserves separate scrutiny. The adapted model described in Chapter 6 allows for that.
- A new conceptualisation of the relationships between agents that participate in an online SNS (Figure 5) was useful in identifying the different interest groups affected by risk.
- Degrees of privacy (Figure 11) based on closeness to the individual is a simplified model of personal data that consolidates the approaches of other workers in the field. This simplified model makes it easier to understand how different measures expose users to risk.

This research is timely and topical. At the time of writing, the new European Data Protection Regulation had been passed by the European Parliament and was in the process of being finalised before being implemented. It is broadly based on data protection principles, but introduces several new provisions such as: the right to be forgotten, the right to transfer personal profiles to new services, and the duty of data controllers to report data breaches. This research provides a unique perspective on regulation and it is hoped that policy makers and legislators will be able to call on some of the concepts, insights and methods developed here as they finalise the legislation and set up mechanisms to enforce and monitor it.

Although the research is UK-based and European in perspective, the research methods used and concepts developed may be applied more widely. Large emerging economies such as that of Brazil, are already heavy users of social media and are increasingly attractive to advertisers trying to reach large, aspirant populations with increasing disposable income. This research may provide a basis for comparison with Europe and perhaps help them to avoid some of the mistakes that have already been made.

Finally, there is a move to greater user control, based on user education and awareness. This research highlights the importance of user attitudes and its significance in managing risks to individuals. Comparisons have been made with road safety campaigns, and this represents a useful adjunct to statutory regulation.

GLOSSARY

API	Application Programming Interface
ASA	Advertising Standards Authority
BCS	British Computer Society
BPR	Best Practice Resource
BRIC	Brazil, Russia, India and China
CAP	Committee on Advertising Practice
CILIP	Chartered Institute of Library and Information Professionals
CMA	Computer Misuse Act 1990
DMA	Direct Marketing Association
DPA	Data Protection Act 1998
DPD	Data Protection Directive (95/46/EC)
DPR	Data Protection Regulation 2012
EC	European Commission
ECHR	European Convention on Human Rights
ECJ	European Court of Justice
EDAA	European Digital Advertising Alliance
EDPS	European Data Protection Supervisor
EESC	European Economic and Social Committee
EP	European Parliament
EU	European Union
EULA	End User Licence Agreement
FTC	Federal Trade Commission
HRA	Human Rights Act 1998
IABUK	Internet Advertisers Bureau UK
ICO	Information Commissioner's Office
ID	Identity or Identifier
IP	Internet Protocol
ISP	Internet Service Provider
LIS	Library and Information Service
NHS	National Health Service
OBA	Online Behavioural Advertising
OECD	Organisation for Economic Cooperation and Development
Ofcom	Office of Communications
OII	Oxford Internet Institute
OSN	Online Social Network (see also SNS)

PbC	Privacy by Design
PECR	Privacy and Electronic Communications Regulation
PPI	Payment Protection Insurance
RIPA	Regulation of Investigatory Powers Act 2000
SI	Statutory Instrument
SNS	Social Networking Service
SRO	Self-Regulatory Organization
UK	United Kingdom of Great Britain and Northern Ireland
URL	Uniform Resource Locator
W3C	Worldwide Web Consortium
WP	Working Party

REFERENCES

LEGISLATION CITED

TREATIES

Convention for the Protection of Human Rights and Fundamental Freedoms [ETS No.5] (Rome, 4 Nov. 1950). Treaty Series No. 071/1953 : Cmd. 8969 UK Ratification: 8 March 1951

Treaty of Accession of Denmark, Ireland and the United Kingdom 1972. *Official Journal L 73*, 27.03.1972

Treaty on European Union 1992 (Maastricht Treaty). *Official Journal C191*, 29.07.1992

Treaty on the Functioning of the European Union 2007. *Official Journal C 83*, 30.3.2010, p56

Universal Declaration of Human Rights (New York, 10 Dec. 1950). United Nations

Vienna Convention on the Law of Treaties 1969. United Nations, Treaty Series vol. 1155, p. 331

UK PRIMARY LEGISLATION, STATUTORY INSTRUMENTS AND BILLS

Communications Act 2003

Computer Misuse Act 1990

Constitutional Reform Act 2005

Consumer Protection Act 1987

Data Protection Act 1998

Draft Communications Data Bill, Cm8359. Norwich: TSO on behalf of the Home Office, 2012

Freedom of Information Act 2000

Human Rights Act 1998

Privacy and Electronic Communications (EC Directive) Regulations. SI 2003/2426

Privacy and Electronic Communications (EC Directive) (Amendment) Regulations. SI 2004/1039

Privacy and Electronic Communications (EC Directive) (Amendment) Regulations. SI 2011/1208

Regulation of Investigatory Powers Act 2000

Transfer of Tribunal Functions Order SI 2010/22

EUROPEAN UNION LEGISLATION

Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (Data Protection Directive) *Official Journal L 281*, 23.11.1995, pp.31-50

Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009 amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services. European Union: *Official Journal L 337*, 18.12.2009, p. 11–36.

Proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). COM (2012) 11 final, 2012/0011 (COD), Brussels, 25 January 2012

Privacy and Electronic Communications Directive (ePrivacy Directive) (2002/58/EC) European Union: *Official Journal L 201 , 31/07/2002, p37-47*

NON-UK LEGISLATION

Children's Online Privacy Protection Act 1998 (COPPA), United States of America

Federal Trade Commission Act 2006, United States of America

Homeland Security Act 2002, United States of America

Marco Civil da Internet 2014. (Lei No. 12,965) Brazil.

CASES CITED

Applause Store Productions Ltd & Anor. v Raphael [2008] EWHC

Google Spain SL, Google Inc. v Agencia Española de Protección de Datos (AEPD), Mario Costeja González [2014] CJEU

J19 v Facebook Ireland [2013] NIQB

Prince Albert v Strange [1849] EWHC

Teggart v Teletech UK Ltd [2012] NIIT

BIBLIOGRAPHY

Agate, J. & Ledward, J., 2013. Social Media: how the net is closing in on cyber bullies. *Entertainment Law Review*, 24(8), pp.263–268.

Aldhouse, F., 2013. Data Protection in Europe – some thoughts on reading the academic manifesto. *Computer Law & Security Review*, 29(3), pp.289–292.

Alexa, 2014. Top 500 Global Sites. Available at: http://www.alexa.com/topsites [Accessed August 20, 2014].

Anderson, J., 2013. *Privacy Engineering for Social Networks*. Cambridge: Computer Laboratory.

Andrienko, G. et al., 2013. Report from Dagstuhl. *ACM SIGMOBILE Mobile Computing and Communications Review*, 17(2), p.7.

Andrienko, G. & Andrienko, N., 2012. Privacy Issues in Geospatial Visual Analytics. In G. Gartner & F. Ortag, eds. *Advances in Location-Based Services. 8th International Conference on Location Based Services, 2011, Vienna*. Heidelberg: Springer, pp. 239–246.

Angwin, J. & Mcginty, T., 2010. Personal Information Exposed via Biggest U.S. Websites - Protect Your Privacy. *Wall Street Journal*. Available at: http://online.wsj.com/article/SB10001424052748703977004575393173432219064.html [Accessed September 16, 2012].

Anon, The social economy: Unlocking value and productivity through social technologies. Available at: http://www.mckinsey.com/insights/high_tech_telecoms_internet/the_social_economy [Accessed April 24, 2015].

Article 29 Data Protection WP, 2013. *Opinion 02/2013 on Apps on Mobile Devices*, Brussels, European Union.

ASA, 2013. *ASA Half-Year Report on the Regulation of Online Behavioural Advertising. February 2013-June 2013*, London.

Auerbach, D., 2013. You Are What You Click. *Nation*, 296(9), pp.27–34.

Aven, T. & Renn, O., 2009. On Risk Defined as an Event Where the Outcome is Uncertain. *Journal of Risk Research*, 12(1), pp.1–11.

Baldwin, R. & Cave, M. eds., 1999. *Understanding Regulation: theory, strategy and practice* 1st ed., Oxford: Oxford University Press.

Baldwin, R., Cave, M. & Lodge, M., 2012. *Understanding Regulation : theory, strategy, and practice* 2nd ed., Oxford: Oxford University Press.

Barth, A., 2011. *HTTP State Management Mechanism. Request for comments 6265*, Berkeley, CA.

Bawden, D. & Robinson, L., 2012. *Introduction to Information Science*, London: Facet Publishing.

Bawden, D. & Robinson, L., 2013. No Such Thing as Society? On the individuality of information behavior. *Journal of the American Society for Information Science & Technology*, 64(12), pp.2587–2590.

BBC News, 2013a. Arrests Made in Brian Holloway's Trashed House Party. *BBC News*. Available at: http://www.bbc.co.uk/news/world-us-canada-24293414 [Accessed February 11, 2014].

BBC News, 2014. Facebook Apologises to Drag Queens Over Real Name Use. Available at: http://www.bbc.co.uk/news/technology-29454375 [Accessed October 9, 2014].

BBC News, 2011. Facebook U-turns on Phone and Address Data Sharing. Available at: http://www.bbc.co.uk/news/technology-12214628 [Accessed September 16, 2012].

BBC News, 2013b. Professor Mary Beard: "Why I Shamed Twitter Troll." Available at: http://www.bbc.co.uk/news/uk-23502792 [Accessed November 19, 2014].

BBC News, 2012. Wikileaks revelations. *BBC News Online*. Available at: http://www.bbc.co.uk/news/world-11863274 [Accessed July 7, 2014].

Beck, U., 1992. *Risk Society : towards a new modernity*, London : Sage.

Bedi, M., 2014. Social Networks, Government Surveillance, and the Fourth Amendment Mosaic Theory. *Boston University Law Review*, 94(6), pp.1809–1880.

Bélanger, F. & Crossler, R.E., 2011. Privacy in the Digital Age: a review of information privacy research in information systems. *MIS Quarterly*, 35(4), pp.1017–1041.

Better Regulation Commission, 2006. *Risk, Responsibility and Regulation: whose risk is it anyway?*, London: Better Regulation Commission.

Blacklock, W.D., 2003. *A Question of Regulation: a study of the regulation of qualifications in England (EdD Thesis)*. University of London Institute of Education.

Blank, G., 2010. Trust and the Internet : 2003-2009. In *OxIS Workshop, London 5 October 2010*. Oxford: Oxford Internet Institute.

Bogdanor, V., 2009. *The New British Constitution*, Oxford: Hart Publishing.

Bond, R., 2010. Data Ownership in Social Networks - a very personal thing. *Privacy and Data Protection*, 11(1), pp.1–5.

Bonneau, J. & Preibusch, S., 2009. The Privacy Jungle: on the market for data protection in social networks. In *Workshop on the Economics of Information Security, 2009*.

Boudreaux, C., 2014. Social Media Policy Database. Available at: http://socialmediagovernance.com/policies/ [Accessed February 4, 2015].

Boyd, D., 2010. The Future of Privacy: how privacy norms can inform regulation. In *32nd International Conference of Data Protection and Privacy Commissioners, Jerusalem, 29 October 2010*. Jerusalem.

Boyd, D., 2012. The Politics of "Real Names". *Communications of the ACM*, 55(8), pp.29–31.

Boyd, D. & Hargittai, E., 2010. Facebook Privacy Settings: who cares? *First Monday*, 15(8), p.2pp.

Boyd, D.M. & Ellison, N.B., 2007. Social Network Sites: definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1), p.11.

Bradwell, P., 2010. *Private Lives: a people's inquiry into personal information*, London: Demos.

Brady, H.E. & Collier, D. eds., 2004. *Rethinking Social Inquiry: diverse tools, shared standards*, Lanham, MD; Oxford: Rowman & Littlefield.

British Standards Institution, 2010. *BS ISO 31000:2009 Risk management — Principles and guidelines*, London.

Brown, I. & Marsden, C.T., 2013. *Regulating Code : good governance and better regulation in the information age*, Cambridge, MA: The MIT Press.

Brunger, J., 2010. Location Based Services: a review of the regulatory framework. *Data Protection Ireland*, 3(5), pp.9–14.

Buckley-Owen, B., Cooke, L. & Matthews, G., 2012. Information Policymaking in the United Kingdom: the role of the information professional. *Journal of Information Policy*, 2.

Butler, E., 2011. Privacy Setting Awareness on Facebook and Its Effect on User-Posted Content. *Human Communication*, 14(1), pp.39–55. 17p. 4 Charts.

Cannataci, J. & Bonnici, J.P.M., 2003. Can Self-regulation Satisfy the Transnational Requisite of Successful Internet Regulation. *International Review of Law, Computers & Technology*, 17(1), pp.51–61.

Carmagnola, F., Osborne, F. & Torre, I., 2014. Escaping the Big Brother: an empirical study on factors influencing identification and information leakage on the Web. *Journal of Information Science*, 40(2), pp.180–197.

Cavazza, F., 2010. The Social Media Landscape 2011. Available at: http://www.fredcavazza.net/2010/12/14/social-media-landscape-2011/ [Accessed September 16, 2012].

Cavoukian, A., 2012. Privacy by Design [Leading Edge]. *IEEE Technology and Society Magazine*, 31(4), pp.18–19.

Charles Booth Online Archive, 2013. Charles Booth and the survey into life and labour in London (1886-1903). *London School of Economics*. Available at: http://booth.lse.ac.uk/ [Accessed February 3, 2015].

Charmaz, K., 2006. *Constructing Grounded Theory : a practical guide through qualitative analysis*, London: Sage.

Christiansen, L., 2011. Personal Privacy and Internet Marketing: an impossible conflict or a marriage made in heaven? *Business Horizons*, 54(6), pp.509–514.

Clapperton, G., 2009. *This is Social Media: tweet, blog, link and post your way to business success*, Chichester, UK: Capstone.

Clarke, R., 1999. Internet Privacy Concerns Confirm the Case for Intervention. *Communications of the ACM*, 42(2), pp.60–67.

CNIL, 2012. *Google Privacy Policy: main findings and recommendations*, Paris: Commission Nationale de l'Informatique et des Libertés.

Collins, R., 2006. Internet Governance in the UK. *Media Culture and Society*, 28(3), pp.337–358.

Comm, J., 2010. *Twitter Power 2.0: how to dominate your market one Tweet at a time*, Hoboken, NJ: John Wiley & Sons.

Committee of Advertising Practice, 2014. *The CAP Code: the UK code of non-broadcast advertising, sales promotion and direct marketing* 12th ed., London: TSO.

Connolly, C., 2008. *The US Safe Harbor - Fact or Fiction? (2008)*, Prymont, NSW, Australia.

Cooke, L., 2004. *Regulating the Internet: policy and practice with reference to the control of Internet access and content (PhD Thesis)*. Loughborough University.

Cooke, L. & Hall, H., 2013. Facets of DREaM: a social network analysis exploring network development in the UK LIS research community. *Journal of Documentation*, 69(6), pp.786–806.

Court of Justice of the European Union, 2014. *Google Spain SL, Google Inc. v. Agencia Española de Protección de Datos (AEPD), Mario Costeja González*,

Cox, C., 2014. Protecting Victims of Cyberstalking, Cyberharassment, and Online Impersonation through Prosecutions and Effective Laws. *Jurimetrics: The Journal of Law, Science & Technology*, 54(3), pp.277–302.

David, M. & Sutton, C.D., 2011. *Social Research : an introduction*, Los Angeles: Sage.

Delbridge, R. & Kirkpatrick, I., 1994. Theory and Practice of Participant Observation. In V. J. Wass & P. E. Wells, eds. *Principles and Practice in Business and Management Research*. Aldershot; Brookfield VT: Dartmouth, pp. 35–62.

Denham, E., 2009. *Report of Findings into the Complaint Filed by the Canadian Internet Policy and Public Interest Clinic (CIPPIC) against Facebook Inc.*, Ottawa.

Desai, MonicaLodge, DeborahGates, MelodiWolvin, MariaLouer, G., 2012. The FTC Privacy Report: What the report means and how you can get ahead of enforcement trends by

implementing best practices now. *International Journal of Mobile Marketing*, 7(2), pp.26–36.

Direct Marketing Association, 2012. Regulators Find Flaws in Data Protection Changes. Available at: http://dma.org.uk/toolkit/regulators-find-flaws-data-protection-changes [Accessed July 18, 2014].

Doctorow, C., 2014. Privacy Technology Everyone Can Use Would Make Us All More Secure. *The Guardian*.

Dry, S., 2007. *Fishermen and Forecasts: how barometers helped make the metrological department safer in Victorian Britain (Discussion Paper No. 46)*, London: LSE Centre for Analysis of Risk and Regulation.

Dutton, W.H. & Blank, G., 2013. Cultures of the Internet: the Internet in Britain. Oxford Internet Survey 2013. Available at: http://oxis.oii.ox.ac.uk/sites/oxis.oii.ox.ac.uk/files/content/files/publications/OxIS_2013.pdf [Accessed May 19, 2014].

Dutton, W.H. & Blank, G., 2011. Next Generation Users: the Internet in Britain. Oxford Internet Survey 2011 Report.

EDAA, 2015. Your Online Choices: a guide to online behavioural advertising. Available at: http://www.youronlinechoices.com/uk/ [Accessed February 23, 2015].

Edwards, J., 2014. This Chart of Historic App Downloads Shows Just How Huge Snapchat is Becoming. *Business Insider*. Available at: http://www.businessinsider.com/snapchat-app-downloads-statistics-and-data-2014-5 [Accessed August 20, 2014].

Ellison, E.B. & Boyd, D.M., 2013. Sociality through Social Network Sites. In W. H. Dutton, ed. *The Oxford Handbook of Internet Studies*. Oxford: Oxford University Press, pp. 151–172.

European Commission, 2010. *A Comprehensive Approach on Personal Data Protection in the European Union (COM/2010/609)*, Brussels, European Union.

European Council, 2010. *The Stockholm Programme — an open and secure Europe serving and protecting citizens*, Official Journal C 115/01, 04/05/2010, p1-38.

European Economic and Social Committee, 2010. *Opinion of the European Economic and Social Committee on the "impact of social networking sites on citizens/consumers" (own-initiative opinion)*, European Union: Official Journal C 128/12, 18 May 2010, p69-73.

European Parliament, 2009. *Strengthening Security and Fundamental Freedoms on the Internet P6_TA(2009)0194*, European Union: Official Journal CE 117/206 6.5.2010.

European Privacy Authorities, 2012. *Letter to Larry Page, Google from the European Privacy Authorities 16 October 2012*, Brussels.

europe-v-facebook.org, 2014. Europe versus Facebook. Available at: http://www.europe-v-facebook.org/ [Accessed February 5, 2015].

Farr, C., 2013. Can You Trust Facebook with Your Genetic Code? *VentureBeat*. Available at: http://venturebeat.com/2013/10/07/can-you-trust-facebook-with-your-genetic-code/ [Accessed October 10, 2013].

Federal Trade Commission, 2010. *Protecting Consumer Privacy in an Era of Rapid Change: a proposed framework for businesses and policymakers*, Washington DC: FTC.

Feldman, D., 1989. The Nature of Legal Scholarship. *The Modern Law Review*, 52(4), pp.498–517.

Fischhoff, B., Watson, S.R. & Hope, C., 1984. Defining Risk. *Policy Sciences*, 17(2), pp.123–139.

Floridi, L., 2014. Google Ethics Adviser: the law needs bold ideas to address the digital age. *The Guardian*.

Garrie, D.B. & Wong, R., 2010. Social Networking: opening the floodgates to "personal data." *Computer and Telecommunications Law Review*, 16(6), pp.167–175.

Get Safe Online Limited, 2015. Get Safe Online. Available at: http://www.getsafeonline.org/ [Accessed February 21, 2015].

Gilbert, F., 2014. Proposed EU Data Protection Regulation - Issues to consider when planning for the future regime. *Internet Law*, 17(12), pp.1, 13–24.

Gomez, J., Pinnick, T. & Soltani, A., 2009. *Know Privacy*, Berkeley CA: University of California Berkeley, School of Information.

Gray, O. & Mills-Wade, A., 2011. *EASA Best Practice Recommendation on Online Behavioural Advertising*, Brussels: European Advertising Standards Alliance.

Gray, T., Zeggane, T. & Maxwell, W., 2008. US and EU Authorities Review Privacy Threats on Social Networking Sites. *Entertainment Law Review*, 19(4), pp.69–74.

Greenwald, G., 2013. NSA Collecting Phone Records of Millions of Verizon Customers Daily. *The Guardian*.

Gregor, S., 2006. The Nature of Theory in Information Systems. *MIS Quarterly*, 30(3), pp.611–642.

Gurses, S., 2014. Can You Engineer Privacy? *Communications of the ACM*, 57(8), pp.20–23.

Habermas, J. & Burger, T., 1989. *The Structural Transformation of the Public Sphere: an inquiry into a category of bourgeois society*, Cambridge: Polity.

Hammock, M.R. & Rubin, P.H., 2011. *Applications Want to be Free: privacy against information*, Washington DC: Technology Policy Institute.

Hans, G.S., 2012. Privacy Policies, Terms of Service and FTC Enforcement: broadening unfairness regulation for a new era. *Michigan Telecommunications & Technology Law Review*, 19(1), pp.120–163.

Hans-Bredow Institut, 2006. *Study on Co-Regulation Measures in the Media Sector. Final Report*, Brussels: Euroepan Commission.

Hansson, S.O., 2004. Fallacies of Risk. *Journal of Risk Research*, 7(3), pp.353–360.

Hansson, S.O., 2010. Risk: objective or subjective, facts or values. *Journal of Risk Research*, 13(2), pp.231–238.

Harvard University, 2015. Berkman Center for Internet and Society. Available at: http://cyber.law.harvard.edu/ [Accessed February 4, 2015].

Hastak, M. & Culnan, M.J., 2010. Online Behavioral Advertising "Icon" Study: summary of key results. *Future of Privacy Forum*, 25.

Haufler, V., 2001. *A Public Role for the Private Sector: industry self-regulation in a global economy*, Washington DC: Carnegie Endowment for International Peace.

Haynes, D., 2012. *Access to Personal Data in Social Networks: measuring the effectiveness of approaches to regulation (MPhil to PhD transfer report)*. City University London.

Haynes, D., 2014a. Forget the Right to be Forgotten, Other Means Exist. *The Conversation*. Available at: http://theconversation.com/forget-the-right-to-be-forgotten-other-means-exist-29117 [Accessed July 21, 2014].

Haynes, D., 2014b. Social Media and Risk. *CILIP Update*, (June 2014), pp.30–32.

Haynes, D., 2011. Social Networks in the Workplace - some data protection issues. *Free Pint*, (1 December 2011).

Haynes, D. & Robinson, L., 2015. Defining User Risk in Social Networking Services. *Aslib Journal of Information Management*, 67(1), pp.94–115.

Haythornthwaite, R., 2006. *The Regulation of Risk : setting the boundaries*, Bath: University of Bath.

Heffernan, S., 2011. UK Financial Reform Post Crisis: Is more regulation the answer? In T. Green, CJ; Pentecost, EJ; WeymanJones, ed. *Financial Crisis and the Regulation of Finance*. Cheltenham: Edward Elgar Publishing Ltd, pp. 193–211.

Helft, M. & Wortham, J., 2010. Facebook Bows to Pressure over Privacy. *New York Times*, p.B1. Available at: http://www.nytimes.com/2010/05/27/technology/27facebook.html [Accessed November 13, 2014].

Heyvaert, V., 2011. Governing Climate Change: towards a new paradigm for risk regulation. *The Modern Law Review*, 74(6), pp.817–844.

Hjørland, B., 2002. Domain Analysis in Information Science: eleven approaches - traditional as well as innovative. *Journal of Documentation*, 58(4), pp.422–462.

Hooper, V. & Kalidas, T., 2012. Acceptable and Unacceptable Behaviour on Social Networking Sites: a study of the behavioural norms of youth on Facebook. *Electronic Journal of Information Systems Evaluation*, 15(3), pp.259–268.

Hustinx, P., 2010. Opinion of the European Data Protection Supervisor on the Current Negotiations by the European Union of an Anti-Counterfeiting Trade Agreement (ACTA). *Official Journal C*, 147/01, pp.1–13.

Hutter, B.M., 2006. Risk, Regulation and Management. In J. Taylor-Gooby, Peter; Zinn, ed. *Risk in Social Science*. Oxford: Oxford University Press, pp. 202–227.

Hutter, B.M., 2005. *The Attractions of Risk-based Regulation: accounting for the emergence of risk ideas in regulation*, London: LSE, Centre for Analysis of Risk and Regulation.

IABUK, 2012. The Evolution of Online Display Advertising (IABUK). Available at: https://www.youtube.com/watch?v=1C0n_9DOlwE.

Information Commissioner's Office, 2012. *Anonymisation: managing data protection risk code of practice*, Wilmslow: ICO.

Information Commissioner's Office, 2014. Information Commissioner's Office - Home Page. Available at: https://ico.org.uk/ [Accessed December 29, 2014].

Information Commissioner's Office, 2008. *Privacy by Design*, Wilmslow: ICO.

Information Commissioner's Office, 2009. *Your Personal Little Book about Protecting your Personal Information*, Wilmslow: ICO.

International Trade Administration, 2009. U.S.-EU Safe Harbor Overview. Available at: http://www.export.gov/safeharbor/eu/eg_main_018475.asp [Accessed February 13, 2015].

Jay, R., 2003. *Data protection : law and practice* A. Hamilton, ed., London: Sweet & Maxwell.

Jimenez, D.L., 2009. Doctrina: Privacidad y Seguridad en el Comercio Electrónico: Nuevos Retos y Desafíos. (Spanish). *Cuadernos de Derecho y Comercio*, (52).

Johnson, B., 2010. Privacy is No Longer a Social Norm says Facebook Founder. *The Guardian*. Available at: http://www.theguardian.com/technology/2010/jan/11/facebook-privacy [Accessed February 6, 2015].

Justice, G., 2010. *ARTICLE 29 DATA PROTECTION WORKING PARTY Opinion 2 / 2010 on online behavioural advertising*, European Union: Article 29 Data Protection Working Party.

Kaplan, A.M. & Haenlein, M., 2010. Users of the World, Unite! The challenges and opportunities of social media. *Business Horizons*, 53(1), pp.59–68.

Kiser, A., 2011. Benefits and Risks of Social Networking Sites: Should they also be Used to Harness Communication in a College or University Setting. *International Journal of Digital Literacy and Digital Competence*, 2(4), pp.1–13.

Kontaxis, G., Polychronakis, M. & Markatos, E., 2012. Minimizing Information Disclosure to Third Parties in Social Login Platforms. *International Journal of Information Security*, 11(5), pp.321–332.

Krishnamurthy, B. & Wills, C.E., 2009. On the Leakage of Personally Identifiable Information via Online Social Networks. In *Proceedings of the 2nd ACM Workshop on Online Social Networks - WOSN '09*. New York NY: ACM Press, p. 7.

Krotoszynski Jr., R.J., 2013. The Polysemy of Privacy. *Indiana Law Journal*, 88(3), pp.881–918.

Külcü, Ö. & Henkoğlu, T., 2014. Privacy in Social Networks: an analysis of Facebook. *International Journal of Information Management*, 34(6), pp.761–769.

Kuzma, J., 2011. Empirical Study of Privacy Issues among Social Networking Sites. *Journal of International Commercial Law & Technology*, 6(2), pp.74–84.

Langheinrich, M. & Karjoth, G., 2010. Social Networking and the Risk to Companies and Institutions. *Information Security Technical Report*, 15(2), pp.51–56.

LaRose, R., Lin, C. & Eastin, M., 2003. Unregulated Internet Usage: addiction, habit, or deficient self-regulation? *Media Psychology*, 5(3), pp.225–253.

Lee, D., 2014. Drag Queens in Facebook Name Row. *BBC News Online*. Available at: http://www.bbc.co.uk/news/technology-29175102 [Accessed February 5, 2015].

Leigh, D. & Harding, L., 2011. *Wikileaks: inside Julian Assange's war on secrecy*, London: Guardian Books.

Lemieux, R., 2012. Fictional Privacy among Facebook Users. *Psychological Reports*, 111(1), pp.289–292.

Lessig, L., 2006. *Code* 2nd ed., New York; London: BasicBooks.

Levin, A. & Abril, P.S., 2009. Two Notions of Privacy Online. *Vanderbilt Journal of Entertainment & Technology Law*, 11(4), pp.1001–1051.

Liu, K. & Terzi, E., 2010. A Framework for Computing the Privacy Scores of Users in Online Social Networks. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 5(1), pp.1–30.

Livingstone, S., 2013. Online risk, harm and vulnerability: Reflections on the evidence base for child Internet safety policy. *Zer*, 18(35), pp.13–28.

Lofstedt, R. et al., 2011. The Changing Nature of Communication and Regulation of Risk in Europe. *Journal of Risk Research*, 14(4), pp.409–429.

Lozada, H.R., Kritz, G.H. & Mintu-Wimsatt, A., 2013. The Challenge of Online Privacy to Global Marketers. *Journal of Marketing Development & Competitiveness*, 7(1), pp.54–62.

Lynskey, O., 2012. *Identifying the Objectives of EU Data Protection Regulation and Justifying its Costs (PhD Thesis)*. University of Cambridge, Lucy Cavendish College.

Lynskey, O., 2011. Track[ing] Changes: an examination of EU regulation of online behavioural advertising through a data protection lens. *European Law Review*, 36(6), pp.874–886.

Macgill, S.M. & Siu, Y.L., 2005. A New Paradigm for Risk Analysis. *Futures*, 37(10), pp.1105–1131.

Malhotra, N.K., Kim, S.S. & Agarwal, J., 2004. Internet Users' Information Privacy Concerns (IUIPC): the construct, the scale, and a causal model. *Information Systems Research*, 15(4), pp.336–355.

Mann, B.L., 2009. Social Networking Websites: a concatenation of impersonation, denigration, sexual agressive solicitation, cyber-bullying or happy snapping videos. *International Journal of Law and Information Technology*, 17(3), pp.252–264.

Mansell, R., 2008. *Communication and Information: Towards a Prospective Research Agenda Report on a Workshop, UNESCO, Paris, 20-21 December 2007*, Paris.

Martin, K., 2012. Diminished or Just Different? A factorial vignette study of privacy as a social contract. *Journal of Business Ethics*, 111(4), pp.519–539.

Mayer-Schönberger, V. & Cukier, K., 2013. *Big Data : a revolution that will transform how we live, work and think*, London: John Murray.

McCullagh, K., 2009. Protecting "Privacy" through Control of "Personal" Data Processing: a flawed approach. *International Review of Law, Computers & Technology*, 23(1-2), pp.13–24.

McDonald, T., 2013. Kids + Facebook = Home Invasion? *Business 2 Community*. Available at: http://www.business2community.com/facebook/kids-facebook-home-invasion-0618724 [Accessed October 10, 2013].

McGoldrick, D., 2013. Developments in the Right to be Forgotten. *Human Rights Law Review*, 13(4), pp.761–776.

McKeon, M., 2010. The Evolution of Privacy on Facebook. Available at: http://mattmckeon.com/facebook-privacy/ [Accessed September 15, 2014].

McNeill, P., 2005. *Research Methods*, London: Routledge.

Mendel, T. et al., 2012. *Global Survey on Internet Privacy and Freedom of Expression (Unesco series on internet freedom)*, Paris: UNESCO Publishing.

Millard, C. & Hon, W.K., 2012. Defining "Personal Data" in E-Social Science. *Information, Communication & Society*, 15(1), pp.66–84.

Miller, D., 2011. *Tales from Facebook*, Cambridge: Polity.

Moran, M., 2005. *Politics and Governance in the UK*, Basingstoke: Palgrave Macmillan.

Morrison, S.R., 2014. The System of Domestic Counterterrorism Law Enforcement. *Stanford Law & Policy Review*, 25(2), pp.341–377.

Muir, A. & Oppenheim, C., 2002. National Information Policy Developments Worldwide IV: copyright, freedom of information and data protection. *Journal of Information Science*, 28(6), pp.467–481.

Narayanan, A. & Shmatikov, V., 2009. De-anonymizing Social Networks. In *2009 30th IEEE Symposium on Security and Privacy*. IEEE, pp. 173–187.

Niglas, K., 2010. The Multidimensional Model of Research Methodology. An integrated set of continua. In A. Tashakkori & C. Teddie, eds. *Sage Handbook of Mixed Methods in Social and Behavoral Research*. Los Angeles ; London: Sage, pp. 215–236.

Nissenbaum, H.F., 2010. *Privacy in Context : technology, policy, and the integrity of social life*, Stanford CA: Stanford Law Books.

Nosko, A. et al., 2012. Examining Priming and Gender as a Means to Reduce Risk in a Social Networking Context: can stories change disclosure and privacy setting use when personal profiles are constructed? *Computers in Human Behavior*, 28(6), pp.2067–2074.

Nunan, D. & Yenicioglu, B., 2013. Informed, Uninformed and Participative Consent in Social Media Research. *International Journal of Market Research*, 55(6), pp.791–808.

O'Brien, D. & Torres, A.M., 2012. Social Networking and Online Privacy: Facebook users' perceptions. *Irish Journal of Management*, 31(2), pp.63–97.

OECD, 1980. *OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*, Paris.

Ofcom, 2009. *How People Assess Online Content and Services*, London: Ofcom.

Ofcom, 2008. *Social Networking: a quantitative and qualitative research report into attitudes, behaviours and use*, London: Ofcom.

Oppenheim, C., 2001. *The Legal and Regulatory Environment for Electronic Information*, Tetbury: Infonortics.

Outhwaite, W. & Turner, S.P. eds., 2007. *The SAGE Handbook of Social Science Methodology*, Los Angeles CA; London: Sage Publications.

PCC, 2012. *Newspaper and Magazine Publishing in the U.K. Editors' Code of Practice*, London: Press Complaints Commission.

Pearsall, J. & Hanks, P. eds., 1999. *The New Oxford Dictionary of English*, Oxford: Oxford University Press.

Pickard, A.J., 2013. *Research Methods in Information* 2nd ed., London: Facet.

Pierson, J. & Heyman, R., 2011. Social Media and Cookies: challenges for online privacy. *info*, 13(6), pp.30–42.

Popper, K.R., 1959. *The Logic of Scientific Discovery*, London: Hutchinson.

Powles, J. & Singh, J., 2014. Academic Commentary: Google Spain. *Cambridge Code*. Available at: http://www.cambridge-code.org/googlespain [Accessed July 21, 2014].

Radaelli, C.M., 2010. Regulating Rule-Making via Impact Assessment. *Governance*, 23(1), pp.89–108.

Ragin, C.C., 1994. *Constructing Social Research : the unity and diversity of method*, Thousand Oaks; London: Pine Forge Press.

Ragin, C.C., 2000. *Fuzzy-set Social Science*, Chicago: University of Chicago Press.

Ragin, C.C. & Becker, H.S. eds., 1992. *What is a Case?: exploring the foundations of social inquiry*, Cambridge: Cambridge University Press.

Reay, I., Dick, S. & Miller, J., 2009. A Large-scale Empirical Study of P3P Privacy Policies. *ACM Transactions on the Web*, 3(2), pp.1–34.

Reidenberg, J.R., 1998. Lex Informatica: the formulation of information policy rules through technology. *Texas Law Review*, 76(3), pp.553–584.

Roberts, L., 2010. Facebook Status Updates are "Burglary Risk." *BBC News Online*. Available at: http://www.bbc.co.uk/news/uk-england-birmingham-12062331 [Accessed July 7, 2014].

Robinson, L., 2009. Information Science: communication chain and domain analysis. *Journal of Documentation*, 65(4), pp.578–591.

Rodrigues, R., 2010. Privacy on Social Networks: norms, markets and natural monopoly. In S. Levmore & M. C. Nussbaum, eds. *The Offensive Internet*. Cambridge, MA: Harvard University Press, pp. 237–256.

Room, S., 2007. *Data Protection and Compliance in Context*, Swindon: British Computer Society.

Rosenblum, D., 2007. What Anyone Can Know: the privacy risks of social networking sites. *IEEE Security & Privacy*, 5(3), pp.40–49.

Rowlands, I., Eisenschitz, T.S. & Bawden, D., 2002. Frame Analysis as a Tool for Understanding Information Policy. *Journal of Information Science*, 28(1), pp.31–38.

Rubinstein, I.S. & Good, N., 2013. Privacy by Design: a counterfactual analysis of Google and Facebook privacy incidents. *Berkeley Technology Law Journal*, 28(2), pp.1333–1413.

Saeri, A.K. et al., 2014. Predicting Facebook Users' Online Privacy Protection: risk, trust, norm focus theory, and the theory of planned behavior. *The Journal of Social Psychology*, 154(4), pp.352–69.

Schmidt, E. & Cohen, J., 2013. *The New Digital Age. Reshaping the future of people, nations and business*, London: John Murray.

Schneier, B., 2010. A Taxonomy of Social Networking Data. *IEEE Security and Privacy*, 8(4), p.88.

Schwartz, P.M., 2013. The EU-U.S. Privacy Collision: a turn to institutions and procedures. *Harvard Law Review*, 126(7), pp.1966–2009.

Sellars, S., 2011. Online privacy : do we have it and do we want it ? A review of the risks and UK case law. *European Intellectual Property Review*, 33(1), pp.9–17.

Serpell, A., 2008. Regulation Impact Assessment. *Economic Papers*, Special Ed(March 2008), pp.41–52.

Shaw, M., McCarthy, C. & Dykeman, K., 2012. The Power of Like Europe: how social marketing works for retail brands. , p.19. Available at: http://www.comscore.com/Insights/Presentations-and-Whitepapers/2012/The-Power-of-Like-Europe-How-Social-Marketing-Works-for-Retail-Brands [Accessed September 16, 2012].

Shehab, M. et al., 2012. Access Control for Online Social Networks Third Party Applications. *Computers & Security*, 31(8), pp.897–911.

Sherman, M., 2012. Advising Clients on Internet Privacy Policies. *GPSolo*, 29(6), pp.48–51.

Sipior, J.C., Ward, B.T. & Mendoza, R.A., 2011. Online Privacy Concerns Associated with Cookies, Flash Cookies, and Web Beacons. *Journal of Internet Commerce*, 10(1), pp.1–16.

Slattery, D. & Nellis, J., 2011. Rethinking the Role of Regulation in the Aftermath of the Global Financial Crisis: the case of the UK. *Panoeconomicus*, 58(3), pp.407–423.

Slavtcheva-Petkova, V., Nash, V.J. & Bulger, M., 2015. Evidence on the extent of harms experienced by children as a result of online risks: implications for policy and research. *Information, Communication and Society*, 18(1), pp.48–62.

Smit, E.G., Van Noort, G. & Voorveld, H.A.M., 2014. Understanding Online Behavioural Advertising: user knowledge, privacy concerns and online coping behaviour in Europe. *Computers in Human Behavior*, 32, pp.15–22.

Smith, H.J., Dinev, T. & Xu, H., 2011. Information Privacy Research: an interdisciplinary review. *MIS Quarterly*, 35(4), pp.989–1015.

Sobel, R., 2007. The HIPAA Paradox: the privacy rule that's not. *Hastings Center Report*, 37(4), pp.40–50.

Solove, D.J., 2011. *Nothing to Hide: the false tradeoff between privacy and security*, New Haven, CT: Yale University Press.

Solove, D.J., 2010. Speech, Privacy and Reputation on the Internet. In S. Levmore & M. C. Nussbaum, eds. *The Offensive Internet*. Cambridge, MA: Harvard University Press, pp. 15–30.

Solove, D.J., 2007. *The Future of Reputation: gossip, rumor, and privacy on the Internet*, New Haven, CT; London: Yale University Press.

Solove, D.J. & Hartzog, W., 2014. The FTC and the New Common Law of Privacy. *Columbia Law Review*, 114(3), pp.583–676.

Solovic, S., 2013. The Social Media Dilemma: managing business benefits and personal risks. *Business 2 Community*. Available at: http://www.business2community.com/social-media/social-media-dilemma-managing-business-benefits-personal-risks-0597574 [Accessed October 10, 2013].

Spinello, R.A., 2006. *Cyberethics: morality and law in cyberspace*, Sudbury MA ; London: Jones and Bartlett Publishers.

Squicciarini, A. et al., 2014. Identifying Hidden Social Circles for Advanced Privacy Configuration. *Computers & Security*, 41(3), pp.40–51.

Staksrud, E. & Livingstone, S., 2009. Children and Online Risk: powerless victims or resourceful participants? *Information, Communication & Society*, 12(3), pp.364–387.

Stone, B., 2009. Facebook Plans Changes to Friend Updates. *New York Times*. Available at: http://bits.blogs.nytimes.com/2009/03/04/facebook-plans-changes-to-friend-updates/ [Accessed February 18, 2015].

Story, L. & Stone, B., 2007. Facebook Retreats on Online Tracking. *New York Times*. Available at: http://www.nytimes.com/2007/11/30/technology/30face.html [Accessed February 4, 2015].

Strauß, S. & Nentwich, M., 2013. Social Network Sites, Privacy and the Blurring Boundary Between Public and Private Spaces. *Science and Public Policy*, 40(6), pp.724–732.

Streitfeld, D. & Perlroth, N., 2012. Instagram Reversal Doesn't Appease Everyone. *New York Times*. Available at: http://www.nytimes.com/2012/12/22/technology/instagram-reversal-doesnt-appease-everyone.html [Accessed February 18, 2015].

Stutzman, F., Capra, R. & Thompson, J., 2011. Factors Mediating Disclosure in Social Network Sites. *Computers in Human Behavior*, 27, pp.590–598.

Svantesson, D.J.B., 2014. The Extraterritoriality of EU Data Privacy Law: its theoretical justification and its practical effect on U.S. businesses. *Stanford Journal of International Law*, 50(1), pp.53–117.

Swedlow, B. et al., 2009. Theorizing and Generalizing about Risk Assessment and Regulation through Comparative Nested Analysis of Representative Cases. *Law & Policy*, 31(2), pp.236–269.

Tavani, H.T., 2000. Privacy and Security. In D. Langford, ed. *Internet Ethics*. Basingstoke: Macmillan, pp. 65–95.

Tene, O. & Polonetsky, J., 2012. Symposium Issue - Privacy in the Age of Big Data: a time for big decisions. *Stanford Law Review Online*, 64(63), pp.63–69.

Thomas, K., Grier, C. & Nicol, D.M., 2010. unFriendly: multi-party privacy risks in social networks. In M. Atallah & N. Hopper, eds. *Privacy Enhancing Technologies. 10th International Symposium. Berlin, 21-23 July 2010*. Berlin; Heidelberg: Springer-Verlag, pp. 236–252.

Thomas, R. & Walport, M., 2008. *Data Sharing Review*, London: Ministry of Justice.

Toch, E., Wang, Y. & Cranor, L.F., 2012. Personalization and privacy: a survey of privacy risks and remedies in personalization-based systems. *User Modeling and User-Adapted Interaction*, 22(1-2), pp.203–220.

Torriti, J., 2007. Impact Assessment in the EU: a tool for better regulation, less regulation or less bad regulation? *Journal of Risk Research*, 10(2), pp.239–276.

Toy, A., 2013. Different Planets or Parallel Universes: old and new paradigms for information privacy. *New Zealand Universities Law Review*, 25(5), pp.938–959.

Tulloch, J., 2006. Everyday Life and Leisure Time. In P. Taylor-Gooby & J. Zinn, eds. *Risk in Social Science*. Oxford: Oxford University Press, pp. 117–139.

Turrow, S., 2004. Bridging the Quantitative and Qualitative Divide. In H. Brady & D. Collier, eds. *Rethinking Social Inquiry: diverse tools, shared standards*. Lanham MD ; Oxford: Rowman & Littlefield, pp. 171–179.

University of Oxford, 2015. Oxford Internet Institute - Research. Available at: http://www.oii.ox.ac.uk/research/ [Accessed February 4, 2015].

Vsauce, 2013. Where Do Deleted Files Go? Available at: https://www.youtube.com/watch?v=G5s4-Kak49o [Accessed September 1, 2014].

Wacks, R., 2010. *Privacy : a very short introduction*, Oxford: Oxford University Press.

Wakefield, J., 2014. Cyberbullies: How best to tackle online abuse? *BBC News Online*. Available at: http://www.bbc.co.uk/news/technology-26121199 [Accessed July 7, 2014].

Warren, S.D. & Brandeis, L.D., 1890. The Right to Privacy. *Harvard Law Review*, 4(5), pp.193–220.

Weber, R.H., 2002. *Regulatory Models for the Online World*, The Hague; London: Kluwer Law International.

Weiss, S., 2008. The Need for a Paradigm Shift in Addressing Privacy Risks in Social Networking Applications. In S. Fischer-Hübner et al., eds. *Proceedings of the Third IFIP WG 9.2, 9.6/11.6, 11.7/FIDIS International Summer School on the Future of Identity in the Information Society, Karlstad University, Sweden, August 4-10, 2007*. Frankfurt: Springer, pp. 161–171.

Westin, A.F., 2003. Social and Political Dimensions of Privacy. *Journal of Social Issues*, 59(2), pp.431–453.

Wikipedia, 2014. List of Social Networking Websites. Available at: http://en.wikipedia.org/wiki/List_of_social_networking_websites [Accessed August 20, 2014].

Wilkes, A., 2011. What does privacy in the workplace really mean in Europe ? Part II. *Privacy and Data Protection*, 11(4), pp.1–4.

Willis, L.E., 2014. Why Not Privacy by Default? *Berkeley Technology Law Journal*, 29(1), pp.61–134.

Wills, C.E. & Zeljkovic, M., 2011. A Personalized Approach to Web Privacy: awareness, attitudes and actions. *Information Management & Computer Security*, 19(1), pp.53–73.

Wilson, R.E., Gosling, S.D. & Graham, L.T., 2012. A Review of Facebook Research in the Social Sciences. *Perspectives on Psychological Science*, 7(3), pp.203–220.

Witte, D.S., 2013. Privacy Deleted: Is it too late to protect our privacy online? *Journal of Internet Law*, 17(7), pp.1–28.

Wollan, R., Smith, N. & Zhou, C., 2011. *The Social Media Management Handbook: everything you need to know to get social media working in your business* R. Wollan & N. Smith, eds., Hoboken NJ; Chichester: Wiley.

Woods, L., 2012. User Generated Content: freedom of expression and the role of media in a digital age. In M. Amos, J. Harrison, & L. Woods, eds. *Freedom of Expression and the Media*. Leiden and Boston: Martinus Nijhoff, pp. 141–168.

Wu, T., 2010. *The Master Switch: the rise and fall of information empires*, New York NY: Knopf.

Xu, F., Michael, K. & Chen, X., 2013. Factors Affecting Privacy Disclosure on Social Network Sites: an integrated model. *Electronic Commerce Research*, 13(2), pp.151–168.

Yassine, A. et al., 2012. Knowledge-Empowered Agent Information System for Privacy Payoff in eCommerce. *Knowledge and Information Systems*, 32(2), pp.445–473.

Yokoyama, M.H. & Sekiguchi, T., 2014. The Use of Social Network Sites in the Workplace: a Case Study in Brazilian Companies. *Brazilian Business Review (English Edition)*, 11(2), pp.87–114.

Zell, A.-M., 2014. Due Protection in the Federal Republic of Germany and the European Union: an unequal playing field. *German Law Journal*, 15(3), pp.461–494.

Zittrain, J., 2008. *The Future of the Internet: and how to stop it*, London: Allen Lane.

APPENDICES

APPENDIX A – SENSITIVITY ANALYSIS OF SEARCHES

Sensitivity searches on EbscoHost conducted on 28th November 2014

The table shows the number of hits for each search and the effect of specifying the names of the SNSs included in this research.

TABLE 23 - SENSITIVITY ANALYSIS OF SEARCH STRATEGIES

Search on 'Privacy' AND 'Social Network*'	OR	OR 'Facebook' OR 'Google'	Percentage increase in hits
	8732	12662	45%
Facebook	10142	12662	
Google	11747	12662	
Badoo	8733	12662	
Hi5	8740	12668	
Instagram	8769	12666	
LinkedIn	8859	12683	
Myspace	8812	12695	
Ning	9164	13043	3%
Snapchat	8741	12664	
Twitter	9630	12958	2%
WhatsApp	8741	12664	

APPENDIX B – INITIAL SURVEY OF ATTITUDES TO RISK

Survey questionnaire administered via SurveyGizmo, in April 2011

The workplace survey was distributed via the following British Computer Society (BCS) special interest groups and other group lists:

- BCS Information Risk Management and Assurance LinkedIn Group
- BCS Information Security Specialist Group on LinkedIn
- BCS Law LinkedIn Group
- BCS Internet LinkedIn Group
- BCS Doctoral Consortium LinkedIn Group
- Data Protection and Security LinkedIn Group
- Information and Records Management Society LinkedIn Group
- JISCMAIL Data-Protection
- JISCMAIL Records-Management-UK

Social Networks and Privacy

INTRODUCTION

Department of Information Science, School of Informatics, City University

Please complete this survey to help identify issues that are of concern to users and potential users of social network services.

For the purposes of this survey, social networks are web-accessible services, in which personal profiles are visible to other users of the service. In other words, individual users provide personal data (such as name, address and occupation) in exchange for access to the social network service and to other users.

This survey is intended to explore the issues associated with access to personal data on social networks. This is part of a research degree at City University London to examine ways in which access to personal data is regulated. The objective of this study is to examine the effectiveness of different methods of regulating access to personal data on social networks.

This survey is directed at users and non-users of social network services. Data gathered in this survey will be consolidated so that individual respondents cannot be identified.

David Haynes, March 2011

1.) DO YOU HAVE AN ACTIVE PROFILE ON A SOCIAL NETWORKING SITE (SUCH AS FACEBOOK, TWITTER OR YOUTUBE)?

Social networks are web-accessible services which require individual participants to put up personal profiles that are visible to other users of the service. Individuals often have to provide personal data (such as name, address and occupation) in exchange for access to the social networking service and to other users.

( ) Yes

( ) No

USE OF SOCIAL NETWORKS

2.) WHICH OF THE FOLLOWING SOCIAL NETWORKING SITES DO YOU USE?

Tick all that apply

[ ] Facebook

[ ] Twitter

[ ] Tagged

[ ] Orkut

[ ] MySpace

[ ] Badoo

[ ] LinkedIn

[ ] Others (please specify)

3.) How often do you use a social networking site?

If you use more than one social networking service, please answer for the **most frequently** used service.

( ) Daily or more often

( ) Once a week or more but less than once a day

( ) Once a month or more but less than once a week

( ) Less often than once a month

4.) What do you consider the main benefits of using social networks?

5.) What do you consider the main risks of putting your personal data on social networks?

6.) What measures or precautions should be taken to protect your personal data on social networking services?

Consider what precautions could be taken by: you; service providers; the industry; national governments; and international regulators.

Non-Users

7.) Is there a particular reason why you do not use social networking services?

Please give details

8.) Was personal privacy a factor in your decision not to use social networking services?

If so, please elaborate

9.) Would anything persuade you to use a social networking site such as Facebook, LinkedIn, Twitter or other social network service?

Please give details

ISSUES - ALL RESPONDENTS

10.) ARE THERE ANY DEVELOPMENTS THAT YOU THINK ARE LIKELY TO AFFECT THE USE OF PERSONAL DATA ON SOCIAL NETWORKS IN THE NEXT 2 YEARS?

For instance, are you aware of new services, technology changes, or forthcoming legislation that might have an effect

11.) HOW EFFECTIVE DO YOU THINK THE DATA PROTECTION ACT IS FOR PROTECTING PERSONAL DATA ON SOCIAL NETWORKS?

The Data Protection Action, 1998 governs the handling and use of personal data collected in the UK, regardless of where it is held. It is based on 8 Data Protection Principles that can be found on the Information Commissioner's website.

12.) DO YOU HAVE ANY FURTHER COMMENTS ABOUT THE ISSUES SURROUNDING PROTECTION OF PERSONAL DATA ON SOCIAL NETWORKS?

Please give details below

FUTURE CONTACT

13.) IF YOU ARE INTERESTED IN THE RESULTS OF THIS SURVEY OR IN PARTICIPATING IN A FOLLOW-UP SURVEY, PLEASE INDICATE BELOW:

Please select all that apply. If you give your e-mail address it will only be used for the purposes you have indicated in this response and will not be passed to a third party.

[ ] I would like to be sent a summary of the results of this survey

[ ] I am interested in participating in a follow-up survey

[ ] I am willing to be interviewed

[ ] My e-mail address is:

Thank You!

David Haynes

David Haynes is studying for a research degree at the Department of Information Science in the School of Informatics at City University, London. He can be contacted at: david.haynes.1@city.ac.uk

---

APPENDIX C – SURVEY OF LIS PROFESSIONALS' ATTITUDES TO SNSS IN THE UK

The Survey was distributed via SurveyGizmo in February 2014. It was publicised via the following forums:

Discussion lists on JISCMAIL

- LIS-LINK
- RECORDSMANAGEMENT-UK
- LIS-PROFESSION
- LIS-LIRG

LinkedIn Groups

- LIS Research Methods
- Information Research
- CILIP on LinkedIn
- Information and Records Management Society Group
- ISKOUK
- London Information and Knowledge Exchange

Twitter

- Personal Twitter feed @jdavidhaynes and #citylis

Social Networks, Risk and Regulation

INTRODUCTION

Hi there!

Thanks for following the link to this City University survey.

This short survey (no more than 15 minutes) seeks your views on the risks associated with online social networking. It specifically looks at the risks to users in the United Kingdom and the ways in which those risks might be managed. The survey is part of a PhD research study to compare different ways of regulating access to personal data gathered by online social networking providers.

Online social networking is based on web-accessible services, which allow users to connect with other users to form social or professional networks. This usually involves setting up a personal profile, which is visible to other users.

In line with City University's research policy, participation in this survey is voluntary. You have the right to withdraw from the survey at any time. Data gathered in this survey will be consolidated so that individual respondents cannot be identified. The data will be used for academic research purposes only. At the end of the survey there will be a consent statement which you will need to confirm before submitting the completed questionnaire.

David Haynes, February 2014

BEFORE WE BEGIN WE NEED TO FIND OUT WHETHER THIS SURVEY IS RELEVANT TO YOU. WHERE IN THE UK DO YOU LIVE?*

For the purposes of this survey, the United Kingdom comprises: England, Wales, Scotland and Northern Ireland. It does not include the Isle of Man or the Channel Islands.

( ) England

( ) Wales

( ) Scotland

( ) Northern Ireland

( ) I do not live in the United Kingdom

[Filter question. Non-UK responses terminated at this point]

USE OF SOCIAL NETWORKS

1) DO YOU HAVE AN ACTIVE PROFILE ON AN ONLINE SOCIAL NETWORKING SERVICE SUCH AS: FACEBOOK, TWITTER OR LINKEDIN?

Online social networks are web-accessible services, which allow users to connect with other users to form social or professional networks. This often means putting up a personal profile that is visible to other users.

( ) Yes

( ) No

2) IF YOU DO USE ONLINE SOCIAL NETWORKING SERVICES, HOW OFTEN DO YOU ACCESS THEM?

Use the blank boxes to add the names of online social networks you regularly use, if they are not included in the list.

	Most days	**Most weeks**	**Occasionally**	**Never**
Facebook	( )	( )	( )	( )
Google+	( )	( )	( )	( )
LinkedIn	( )	( )	( )	( )
Twitter	( )	( )	( )	( )

RISKS

An earlier survey identified a number of risks associated with use of online social networks. This has been followed up by an extensive literature survey. In this section we have identified the main risks reported so far. We would like your views on what you consider to be the most important risks.

For the purposes of this survey **risk** is defined as: "an event of unknown probability that has an adverse effect or consequence".

3) THINKING ABOUT YOUR OWN USE OF ONLINE SOCIAL NETWORKS, HOW CONCERNED ARE YOU PERSONALLY ABOUT THE FOLLOWING RISKS?

Please rank them, with the most important risk at the top.

Please note that this feature is not compatible with early versions of some browsers. If you have difficulty, you can list the risks in the response area for Q4 (the next question).

_____Cyber-bullying or harassment (including stalking)

_____Victim of fraud

_____Identity theft

_____Targeting by official bodies or security agencies

_____Targeting by advertisers

_____Targeting by criminals (e.g. so that they can burgle your home while you are away)

_____Discrimination by employer or potential employer

_____Friends, family or colleagues able to see sensitive personal details

_____Strangers able to see sensitive personal details

_____Physical violence or kidnapping

_____Extortion or blackmail

_____Prosecution by authorities because of crime allegations

4) ARE THERE ANY OTHER RISKS ASSOCIATED WITH YOUR PERSONAL DATA ON ONLINE SOCIAL NETWORKS THAT HAVE NOT BEEN INCLUDED IN THE ABOVE LIST?

_____

MEASURES TO MANAGE RISK

5) WHO YOU THINK SHOULD HAVE PRIMARY RESPONSIBILITY FOR PROTECTING YOUR PERSONAL DATA ON ONLINE SOCIAL NETWORKS?

( ) **Government** (UK or European Union for instance)

( ) **Online social network providers** (e.g. Facebook, Twitter, LinkedIn, Google)

( ) **Advertisers** (who obtain profile data from online social network providers)

( ) **Users**

( ) **Other** (Please specify): _____

6) TO WHAT EXTENT DO YOU AGREE OR DISAGREE WITH THE FOLLOWING STATEMENTS?

These statements all refer to data about you, which is held by social networking services (SNSs) such as Facebook, Twitter or LinkedIn. We are interested in your views about who should be responsible for protecting your personal data.

	Strongly disagree	**Disagree**	**Agree**	**Strongly agree**
Current data protection **legislation is effective** for protecting my personal data	( )	( )	( )	( )
The **SNS providers should be responsible** for protecting my personal data without government interference	( )	( )	( )	( )
The **SNS providers should work with government** to protect my personal data	( )	( )	( )	( )
SNSs should be set up with **maximum privacy as the default** setting	( )	( )	( )	( )
My personal **profile should only be visible to those people or groups that I specify**	( )	( )	( )	( )
SNSs should be **designed with protection of personal data in mind**	( )	( )	( )	( )
There should be **no external regulation** of personal data on SNSs	( )	( )	( )	( )
As a user **I should be responsible** for my own online privacy	( )	( )	( )	( )

7) ARE THERE ANY FURTHER MEASURES THAT YOU THINK SHOULD BE IN PLACE TO PROTECT PERSONAL DATA GATHERED BY ONLINE SOCIAL NETWORKS?

Please give details below.

_____

BACKGROUND INFORMATION

Finally, to help us put the results of this survey into context, could you please answer the following quick questions:

8) WHICH AGE RANGE DO YOU FALL INTO?

( ) under 18

( ) 18-24

( ) 25-34

( ) 35-44

( ) 45-54

( ) 55-64

( ) 65+

9) GENDER

( ) Male

( ) Female

10) ARE YOU A MEMBER OF THE LIS PROFESSION (THIS INCLUDES: LIBRARIANS, INFORMATION SCIENTISTS, KNOWLEDGE MANAGERS, RECORDS MANAGERS, INFORMATION MANAGERS, AND ARCHIVISTS)?

Although this survey is primarily targeted at LIS professionals (including students), the results from all respondents will be included in the final analysis.

( ) Yes

( ) No

CONSENT FORM

In order to complete this survey we need your informed consent to store and process the data provided in your response. If you agree to your response being used, please answer 'Yes' to the question below. If you choose not to proceed, your response will be discarded.

I AGREE THAT MY RESPONSE TO THIS SURVEY CAN BE USED FOR ACADEMIC RESEARCH AND RETAINED FOR FUTURE ACADEMIC STUDY. MY RESPONSE WILL BE AGGREGATED SO THAT MY IDENTITY IS NOT REVEALED IN ANY PUBLICATION OF RESULTS.*

( ) Yes

( ) No

CONSENT CHECK

WOULD YOU LIKE TO RETURN TO THE SURVEY CONSENT FORM? IF YOU CLICK ON 'NO' THIS WILL CONFIRM THAT YOU DO NOT WISH TO PARTICIPATE IN THE SURVEY AND YOUR RESPONSE WILL BE DISCARDED.*

( ) Yes

( ) No

FUTURE CONTACT

IF YOU ARE INTERESTED IN THE RESULTS OF THIS SURVEY OR IN PARTICIPATING IN A FOLLOW-UP STUDY, PLEASE SELECT THE BOX(ES) BELOW:

[ ] I would like to be sent a summary of the results of this survey

[ ] I would be interested in participating in a follow-up study

MY E-MAIL ADDRESS IS:

If you give your e-mail address it will only be used for the purposes you have indicated in this response and will not be passed on to a third party.

Thank You!

Thank you for completing this survey.

David Haynes

David Haynes is currently researching the relationship between risk and regulation of social networking services as part of his PhD studies at the Centre for Information Science at City University London. He can be contacted at: david.haynes.1@city.ac.uk

APPENDIX D – INTERVIEW QUESTIONS

The following individuals were interviewed in the period March-April 2014:

Organisation	Contact name	Date	Type of interview
Bigbrotherwatch	Nick Pickles, Director	17 Apr 2014	Face-to-face
British Computer Society	Peter Harris, Chair, Information Privacy Expert Panel	3 Mar 2014	Telephone
Chartered Institute of Public Relations	Martin Horrox, Regulatory Consultant	2 Apr2014	Face-to-face
CILIP	Guy Daines, Head of Policy	6 Mar 2014	Face-to-face
Committee of Advertising Practice	Malcolm Phillips, Regulatory Policy Manager	6 Mar 2014	Face-to-face
Direct Marketing Association	Chris Combemale, CEO	3 Apr 2014	Face-to-face
Enterprise Privacy Group	Toby Stevens	1 Apr 2014	Face-to-face
Information Commissioner's Office	Ian Bourne	19 Mar 2014	Face-to-face
Information Privacy Group	Toby Stevens	1 Apr 2014	Face-to-face
Internet Advertising Bureau, UK	Nick Stringer, Director, Regulatory Affairs	11 Mar 2014	Face-to-face
JanRain	Russell Loarridge	8 Apr 2014	Telephone
London School of Economics	Orla Lynsky, Assistant Professor of Law	9 Apr 2014	Face-to-face
Mydex	David Alexander, CEO	16 Apr 2014	Face-to-face
Privacy International	Anna Fielder (Chair)	7 Apr 2014	Telephone
tScheme	Richard Trevorah	7 Apr 2014	Telephone
UK Government Department	Anonymous Civil Servant	10 Apr 2014	Face-to-face

INTRODUCTION (FOR PARTICIPANTS)

The purpose of this interview is to find out the views of respondents on the different ways in which access to personal data on social media is regulated in the UK. The study is part of a PhD research project exploring the relationship between personal risk and regulation of online social networking services and follows on from an online survey on risk perceptions among users.

This interview will consist of a series open questions covering:

- Current measures in place for protecting personal data on online social networking services
- Your views on the effectiveness of current measures
- Specific issues and problems associated with personal data
- Potential future measures, including proposed legislation

The interview is expected to last between 45 minutes and one hour in total. It will be recorded (audio recording) so that the notes can be accurately transcribed and analysed. We may need to contact you subsequently for clarification of any points arising from the interview. We will only do this with your permission.

This study is subject to the approval of the City University London, School of Informatics Research Ethics Committee. It is one of the University's requirements that all survey respondents should have consented to participation in the study before participating. [Make sure that the respondent has had a chance to read the participant guidelines and has signed the consent form before beginning the interview.]

QUESTIONS

PRELIMINARIES

Interview with [Name] of [Organisation] on [Date]

BACKGROUND ABOUT YOUR REGULATORY ROLE

Can you please explain your organisation's role in regulating access to personal data?

[Prompts]

- Code of practice – Training – Awareness
- Promotion of good practice
- Target audiences

- Written guidelines
- Current or due for update?

VIEWS ON RISK

What do you think are the main risks that users are exposed to when they use online social networking services? [Prompt with a list of risks identified in previous surveys, if necessary]

Do you think that regulation reduces risks to individual users of social networks? In what ways?

REGULATORY MEASURES IN PLACE

What other measures that you are aware of are in place to protect social media users against misuse of their personal data?

[Service providers and advertisers] What measures does your organisation/industry take (have in place) to protect users against misuse of their data?

[SNS Providers and Advertisers] Do you subscribe to an industry code of practice?

- Who issues the code of practice? Do you have a contact?
- Do you have a copy of the code of practice that I can have?

VIEWS ON LEGISLATION

What is your view of the current Data Protection Act as a way of protecting people against misuse of personal data that they put up on online Social networking services?

- Do you think that the current legislation is effective?
- Do you think that it could be improved? If so, in what ways?

In your view, is this area over- or under-regulated? Why?

Are you familiar with the proposed European Data Protection Regulation currently under discussion?

- Do you think this is an improvement on the current legislation? Why?

REGULATORY EFFECTIVENESS

[Regulators and self-regulators] How do you assess regulatory effectiveness?

[Regulators and self-regulators] Do you think that risk could be used as a way of measuring regulatory effectiveness?

RESPONSIBILITY FOR REGULATION

Who should have primary responsibility for protecting users against misuse of personal data that they put up on SNS profiles? [Prompt: users themselves, industry bodies, system designers, SNS providers, the government, others?]

Why?

FOLLOW-UP

May we quote you?

May we attribute any interview comments to you?

May we identify your organisation?

May we approach you again if any points need clarifying or if we need to follow up any aspect of this interview?

Can you suggest other people or organisations that you think should be consulted as part of this study? – Can I mention your name?

Thank you

APPENDIX E – CASE STUDY PROTOCOL

METHOD

The protocol was developed from an initial survey during the early stages of the research and was supplemented by reviewing the literature. This was followed by a systematic review of online activity and discussions groups about privacy in online social networks. The majority of this investigation took place during the period October 2013 to February 2014 with validation of results in December 2014. The following method was adopted:

BACKGROUND SEARCH

Search on Google search engine for the following: "privacy" and "social media". The top listings were manually scanned to select relevant sites. Candidate sites were visited to assess their relevance to this study. The main criteria were as follows:

- Up to date
- Articles, blogs, postings on privacy in the context of social networking services
- Open and available to general Internet Users (i.e. no membership registration required)

The following were identified as being of particular interest and were tagged using the Delicious.com bookmarking service. They were monitored throughout this study:

Access	https://www.accessnow.org
Digital Civil Rights in Europe	https://edri.org
eMarketer	http://www.emarketer.com
EPIC - Electronic Privacy Information Center	https://epic.org
EUROPE versus FACEBOOK	http://www.europe-v-facebook.org
Future of Privacy Forum	http://www.futureofprivacy.org
GigaOM	https://gigaom.com
La Quadrature du Net	http://www.laquadrature.net
Open Rights Group	http://www.openrightsgroup.org
Privacy by Design	https://privacybydesign.ca
Public Knowledge	https://www.publicknowledge.org
TechCrunch	http://techcrunch.com
Technorati	http://technorati.com
The Guardian	http://www.theguardian.com
Truthdig	http://www.truthdig.com

World Privacy Forum	http://www.worldprivacyforum.org
Your Online Choices	http://www.youronlinechoices.eu

Relevant discussion groups on the following social networking sites were also reviewed:

- Facebook
- LinkedIn
- Twitter
- Google+

TRACKING REGISTRATION

The tests were used to generate data about interactions between users, and SNS systems during the registration process. This approach was used to complement the declared policies of the SNS providers and the literature about the nature of the risks faced by users. The registration tracking activity was intended to independently validate the obligatory data that SNSs say they require in their privacy policies.

1. Log onto each service to determine the minimum registration details to provide insight into exposure of users' personal data on SNSs.
2. Screen capture using the Snipping Tool bundled with Windows.
3. Analyse and tabulate results

APPENDIX F – DATA SETS ON CD-ROM

Results of the 2011 Survey (from Appendix B)

Results of the 2014 Survey (from Appendix C)

Interview Transcripts and Recordings (from Appendix D)

Screenshots from Case Studies (from Appendix E)

Privacy Policies of SNS Providers Downloaded as PDFs

NVivo10 Project File for Content Analysis